

Mathematics and Its Applications

E.I. Gordon, A.G. Kusraev
and S.S. Kutateladze

Infinitesimal Analysis



Kluwer Academic Publishers

Contents

Foreword	vii
Chapter 1. Excursus into the History of Calculus	1
§ 1.1. G. W. Leibniz and I. Newton	2
§ 1.2. L. Euler	5
§ 1.3. G. Berkeley	5
§ 1.4. J. D’Alembert and L. Carnot	6
§ 1.5. B. Bolzano, A. Cauchy, and K. Weierstrass	7
§ 1.6. N. N. Luzin	7
§ 1.7. A. Robinson	9
Chapter 2. Naive Foundations of Infinitesimal Analysis	10
§ 2.1. The Concept of Set in Infinitesimal Analysis	10
§ 2.2. Preliminaries on Standard and Nonstandard Reals	16
§ 2.3. Basics of Calculus on the Real Axis	23
Chapter 3. Set-Theoretic Formalisms of Infinitesimal Analysis	35
§ 3.1. The Language of Set Theory	37
§ 3.2. Zermelo–Fraenkel Set Theory	47
§ 3.3. Nelson Internal Set Theory	64
§ 3.4. External Set Theories	72

§ 3.5. Credenda of Infinitesimal Analysis	80
§ 3.6. Von Neumann–Gödel–Bernays Theory	85
§ 3.7. Nonstandard Class Theory	94
§ 3.8. Consistency of NCT	101
§ 3.9. Relative Internal Set Theory	106
Chapter 4. Monads in General Topology	116
§ 4.1. Monads and Filters	116
§ 4.2. Monads and Topological Spaces	123
§ 4.3. Nearstandardness and Compactness	126
§ 4.4. Infinite Proximity in Uniform Space	129
§ 4.5. Prenearstandardness, Compactness, and Total Boundedness ..	133
§ 4.6. Relative Monads	140
§ 4.7. Compactness and Subcontinuity	148
§ 4.8. Cyclic and Extensional Filters	151
§ 4.9. Essential and Proideal Points of Cyclic Monads	156
§ 4.10. Descending Compact and Precompact Spaces	159
§ 4.11. Proultrafilters and Extensional Filters	160
Chapter 5. Infinitesimals and Subdifferentials	166
§ 5.1. Vector Topology	166
§ 5.2. Classical Approximating and Regularizing Cones	170
§ 5.3. Kuratowski and Rockafellar Limits	180
§ 5.4. Approximation Given a Set of Infinitesimals	189
§ 5.5. Approximation to Composites	199
§ 5.6. Infinitesimal Subdifferentials	204
§ 5.7. Infinitesimal Optimality	219

Chapter 6. Technique of Hyperapproximation	223
§ 6.1. Nonstandard Hulls	224
§ 6.2. Discrete Approximation in Banach Space	233
§ 6.3. Loeb Measure	242
§ 6.4. Hyperapproximation of Measure Space	252
§ 6.5. Hyperapproximation of Integral Operators	262
§ 6.6. Pseudointegral Operators and Random Loeb Measures	272
Chapter 7. Infinitesimals in Harmonic Analysis	281
§ 7.1. Hyperapproximation of the Fourier Transform on the Reals	281
§ 7.2. A Nonstandard Hull of a Hyperfinite Group	294
§ 7.3. The Case of a Compact Nonstandard Hull	307
§ 7.4. Hyperapproximation of Locally Compact Abelian Groups	316
§ 7.5. Examples of Hyperapproximation	327
§ 7.6. Discrete Approximation of Function Spaces on a Locally Compact Abelian Group	340
§ 7.7. Hyperapproximation of Pseudodifferential Operators	355
Chapter 8. Exercises and Unsolved Problems	367
§ 8.1. Nonstandard Hulls and Loeb Measures	367
§ 8.2. Hyperapproximation and Spectral Theory	369
§ 8.3. Combining Nonstandard Methods	371
§ 8.4. Convex Analysis and Extremal Problems	374
§ 8.5. Miscellany	376
Appendix	380
References	385
Notation Index	414
Subject Index	417

Foreword

Nonstandard methods of analysis consist generally in comparative study of two interpretations of a mathematical claim or construction given as a formal symbolic expression by means of two different set-theoretic models: one, a “standard” model and the other, a “nonstandard” model.

The second half of the 20th century is a period of significant progress in these methods and their rapid development in a few directions.

The first of the latter appears often under the name minted by its inventor, A. Robinson. This memorable term *nonstandard analysis* often swaps places with its synonymous versions like *robinsonian* or *classical nonstandard analysis*, remaining slightly presumptuous and defiant.

The characteristic feature of robinsonian analysis is a frequent usage of many controversial concepts appealing to the actual infinitely small and infinitely large quantities that have resided happily in natural sciences from ancient times but were strictly forbidden in mathematics for many decades of the 20th century. The present-day achievements revive the forgotten term *infinitesimal analysis* which reminds us expressively of the heroic bygones of the *Calculus*.

Infinitesimal analysis expands rapidly, bringing about radical reconsideration of the general conceptual system of mathematics. The principal reasons for this progress are twofold. Firstly, infinitesimal analysis provides us with a novel understanding for the method of indivisibles rooted deeply in the mathematical classics. Secondly, it synthesizes both classical approaches to differential and integral calculus which belong to the noble inventors of the latter. Infinitesimal analysis finds newer and newest applications and merges into every section of contemporary mathematics. Sweeping changes are on the march in nonsmooth analysis, measure theory, probability, the qualitative theory of differential equations, and mathematical economics.

The second direction, *Boolean valued analysis*, distinguishes itself by ample usage of such terms as the technique of ascending and descending, cyclic envelopes and mixings, B -sets and representation of objects in $\mathbb{V}^{(B)}$. Boolean valued analysis originated with the famous works by P. J. Cohen on the continuum hypothesis.

Progress in this direction has evoked radically new ideas and results in many sections of functional analysis. Among them we list Kantorovich space theory, the theory of von Neumann algebras, convex analysis, and the theory of vector measures.

The book [1], printed by the Siberian Division of the Nauka Publishers in 1990 and translated into English by Kluwer Academic Publishers in 1994 (see [2]), gave a first unified treatment of the two disciplines forming the core of the present-day nonstandard methods of analysis.

The reader's interest as well as successful research into the field assigns us the task of updating the book and surveying the state of the art. Implementation of the task has shown soon that it is impossible to compile new topics and results in a single book. Therefore, the Sobolev Institute Press decided to launch the series *Nonstandard Methods of Analysis* which will consist of monographs on various aspects of this direction in mathematical research.

The series started with the book [3] whose English edition [4] appeared quite simultaneously. The second in the series was the collection [5] and its English counterpart [6]. This book continues the series and addresses infinitesimal analysis.

The antique treasure-trove contains the idea of an infinitesimal or an infinitely small quantity. Infinitesimals have proliferated for two millennia, enchanting scientists and philosophers but always raising controversy and sometimes despise. After about half a century of willful neglect, contemporary mathematics starts paying rapt attention to infinitesimals and related topics.

Infinitely large and infinitely small numbers, alongside the atoms of mathematics, "indivisibles" or "monads," resurrect in various publications, becoming part and parcel of everyday mathematical practice. A turning point in the evolution of infinitesimal concepts is associated with an outstanding achievement of A. Robinson, the creation of nonstandard analysis now called *robinsonian* and *infinitesimal*.

Robinsonian analysis was ranked long enough as a rather sophisticated, if not exotic, logical technique for corroborating the possibility of use of actual infinities and infinitesimals. This technique has also been evaluated as hardly applicable and never involving any significant reconsideration of the state-of-the-art.

By the end of the 1970s, the views of the place and role of infinitesimal analysis had been drastically changed and enriched after publication of the so-called internal set theory IST by E. Nelson and the external set theories propounded soon after IST by K. Hrbáček and T. Kawai.

In the light of the new discoveries it became possible to consider nonstandard elements as indispensable members of all routine mathematical objects rather than some "imaginary, ideal, or surd entities" we attach to conventional sets by ad hoc reasons of formal convenience.

This has given rise to a new doctrine claiming that every set is either standard or nonstandard. Moreover, the standard sets constitute some frame of reference

“dense” everywhere in the universe of all objects of set-theoretic mathematics, which guarantees healthy translation of mathematical facts from the collection of standard sets to the whole universe.

At the same time many familiar objects of infinitesimal analysis turn out to be “cantorian” sets falling beyond any of the canonical universes in ample supply by formal set theories. Among these “external” sets we list the monads of filters, the standard part operations on numbers and vectors, the limited parts of spaces, etc.

The von Neumann universe fails to exhaust the world of classical mathematics: this motto is one of the most obvious consequences of the new stances of mathematics. Therefore, the traditional views of robinsonian analysis begin to undergo revision, requiring reconsideration of its backgrounds.

The crucial advantage of new ways to infinitesimals is the opportunity to pursue an axiomatic approach which makes it possible to master the apparatus of the modern infinitesimal analysis without learning prerequisites such as the technique and toolbox of ultraproducts, Boolean valued models, etc.

The suggested axioms are very simple to apply, while admitting comprehensible motivation at the semantic level within the framework of the “naive” set-theoretic stance current in analysis. At the same time, they essentially broaden the range of mathematical objects, open up possibilities of developing a new formal apparatus, and enable us to diminish significantly the existent dangerous gaps between the ideas, methodological credenda, and levels of rigor that are in common parlance in mathematics and its applications to the natural and social sciences.

In other words, the axiomatic set-theoretic foundation of infinitesimal analysis has a tremendous significance for science in general.

In 1947 K. Gödel wrote: “There might exist axioms so abundant in their verifiable consequences, shedding so much light upon the whole discipline and furnishing such powerful methods for solving given problems (and even solving them, as far as that is possible, in a constructivistic way), that quite irrespective of their intrinsic necessity they would have to be assumed at least in the same sense as any well established physical theory” [129, p. 521]. This prediction of K. Gödel turns out to be a prophecy.

The purpose of this book is to make new roads to infinitesimal analysis more accessible. To this end, we start with presenting the semantic qualitative views of standard and nonstandard objects as well as the relevant apparatus at the “naive” level of rigor which is absolutely sufficient for effective applications without appealing to any logical formalism.

We then give a concise reference material pertaining to the modern axiomatic expositions of infinitesimal analysis within the classical cantorian doctrine. We have found it appropriate to allot plenty of room to the ideological and historical facets of our topic, which has determined the plan and style of exposition.

Chapters 1 and 2 contain the historical signposts alongside the qualitative motivation of the principles of infinitesimal analysis and discussion of their simplest implications for differential and integral calculus. This lays the “naive” foundation of infinitesimal analysis. Formal details of the corresponding apparatus of nonstandard set theory are given in Chapter 3.

The following remarkable words by N. N. Luzin contains a weighty argument in favor of some concentricity of exposition:

“Mathematical analysis is a science far from the state of ultimate completion with unbending and immutable principles we are only left to apply, despite common inclination to view it so. Mathematical analysis differs in no way from any other science, having its own flux of ideas which is not only translational but also rotational, returning every now and then to various groups of former ideas and shedding new light on them” [335, p. 389].

Chapters 4 and 5 set forth the infinitesimal methods of general topology and subdifferential calculus.

Chapter 6 addresses the problem of approximating infinite-dimensional Banach spaces and operators between them by finite-dimensional spaces and finite-rank operators. Naturally, some infinitely large number plays the role of the dimension or such an approximate space.

The next of kin is Chapter 7 which provides the details of the nonstandard technique for “hyperapproximation” of locally compact abelian groups and Fourier transforms over them.

The choice of these topics from the variety of recent applications of infinitesimal analysis is basically due to the personal preferences of the authors.

Chapter 8 closes exposition, collecting some exercises for drill and better understanding as well as a few open questions whose complexity varies from nil to infinity.

We cannot bear residing in the two-element Boolean algebra and indulge occasionally in playing with general Boolean valued models of set theory. For the reader’s convenience we give preliminaries to these models in the Appendix.

This book is in part intended to submit the authors’ report about the problems we were deeply engrossed in during the last quarter of the 20th century. We happily recall the ups and downs of our joint venture full of inspiration and friendliness. It seems appropriate to list the latter among the pleasant manifestations and consequences of the nonstandard methods of analysis.

E. Gordon
A. Kusraev
S. Kutateladze

References

1. Kusraev A. G. and Kutateladze S. S., Nonstandard Methods of Analysis [in Russian], Nauka, Novosibirsk (1990).
2. Kusraev A. G. and Kutateladze S. S., Nonstandard Methods of Analysis, Kluwer Academic Publishers, Dordrecht etc. (1994).
3. Kusraev A. G. and Kutateladze S. S., Boolean Valued Analysis [in Russian], Sobolev Institute Press, Novosibirsk (1999).
4. Kusraev A. G. and Kutateladze S. S., Boolean Valued Analysis, Kluwer Academic Publishers, Dordrecht etc. (1999).
5. Gutman A. E. et al., Nonstandard Analysis and Vector Lattices [in Russian], Sobolev Institute Press, Novosibirsk (1999).
6. Kutateladze S. S. (ed.), Nonstandard Analysis and Vector Lattices, Kluwer Academic Publishers, Dordrecht etc. (2000).

Chapter 1

Excursus into the History of Calculus

The ideas of differential and integral calculus are traceable from the remote ages, intertwining tightly with the most fundamental mathematical concepts.

We admit readily that to present the evolution of views of mathematical objects and the history of the processes of calculation and measurement which gave an impetus to the modern theory of infinitesimals requires the Herculean efforts far beyond the authors' abilities and intentions.

The matter is significantly aggravated by the fact that the history of mathematics has always fallen victim to the notorious incessant attempts at providing an apologia for all stylish brand-new conceptions and misconceptions. In particular, many available expositions of the evolution of calculus could hardly be praised as complete, fair, and unbiased. One-sided views of the nature of the differential and the integral, hypertrophy of the role of the limit and neglect of the infinitesimal have been spread so widely in the recent decades that we cannot ignore their existence.

It has become a truism to say that “the genuine foundations of analysis have for a long time been surrounded with mystery as a result of unwillingness to admit that the notion of limit enjoys an exclusive right to be the source of new methods” (cf. [65]). However, Pontryagin was right to remark: “In a historical sense, integral and differential calculus had already been among the established areas of mathematics long before the theory of limits. The latter originated as superstructure over an existent theory. Many physicists opine that the so-called rigorous definitions of derivative and integral are in no way necessary for satisfactory comprehension of differential and integral calculus. I share this viewpoint” [401, pp. 64–65].

Considering the above, we find it worthwhile to brief the reader about some turning points and crucial ideas in the evolution of analysis as expressed in the words of classics. The choice of the corresponding fragments is doomed to be subjective. We nevertheless hope that our selection will be sufficient for the reader to acquire a critical attitude to incomplete and misleading delineations of the evolution of infinitesimal methods.

1.1. G. W. Leibniz and I. Newton

The ancient name for differential and integral calculus is “infinitesimal analysis.”[†]

The first textbook on this subject was published as far back as 1696 under the title *Analyse des infiniment petits pour l’intelligence des lignes courbe*. The textbook was compiled by de l’Hôpital as a result of his contacts with J. Bernoulli (senior), one of the most famous disciples of Leibniz.

The history of creation of mathematical analysis, the scientific legacy of its founders and their personal relations have been studied in full detail and even scrutinized. Each fair attempt is welcome at reconstructing the train of thought of the men of genius and elucidating the ways to new knowledge and keen vision. We must however bear in mind the principal differences between draft papers and notes, personal letters to colleagues, and the articles written especially for publication. It is therefore reasonable to look at the “official” presentation of Leibniz’s and Newton’s views of infinitesimals.

The first publication on differential calculus was Leibniz’s article “Nova methodus pro maximis et minimis, itemque tangentibus, quae nec fractals nec irrationales quantitates moratur, et singulare pro illis calculi genus” (see [311]). This article was published in the Leipzig journal “Acta Eruditorum” more than three centuries ago in 1684.

Leibniz gave the following definition of differential. Considering a curve YY and a tangent at a fixed point Y on the curve which corresponds to a coordinate X on the axis AX and denoting by D the intersection point of the tangent and axis, Leibniz wrote: “Now some straight line selected arbitrarily is called dx and another line whose ratio to dx is the same as of $\dots y \dots$ to XD is called $\dots dy$ or difference (*differentia*) \dots of $y \dots$ ”

The essential details of the picture accompanying this text are reproduced in Fig. 1.

By Leibniz, given an arbitrary dx and considering the function $x \mapsto y(x)$ at a point x , we obtain

$$dy := \frac{YX}{XD} dx.$$

In other words, the differential of a function is defined as the appropriate linear mapping in the manner fully acceptable to the majority of the today’s teachers of analysis.

Leibniz was a deep thinker and polymath who believed that “the invention of the syllogistic form ranks among the most beautiful and even the most important

[†] This term was used in 1748 by Leonhard Euler in *Introductio in Analysin Infinitorum* [109] (cf. [239, p. 324]).

discoveries of the human mind. This is a sort of *universal mathematics* whose significance has not yet been completely comprehended. It can be said to incarnate the art of faultlessness ...” [313, pp. 492–493]. Leibniz understood definitely that the description and substantiation of the algorithm of differential calculus (in that way he referred to the rules of differentiation) required clarifying the concept of tangent. He proceeded with explaining that “we have only to keep in mind that to find a tangent means to draw the line that connects two points of the curve at an infinitely small distance, or the continued side of a polygon with an infinite number of angles which for us takes the place of the curve.” We may conclude that Leibniz rested his calculus on appealing to the structure of a curve “in the small.”

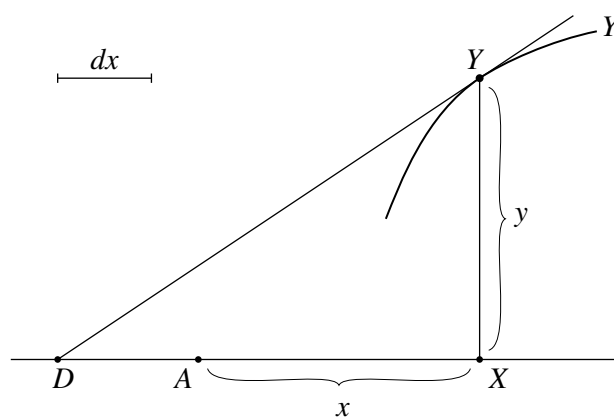


Fig. 1

At that time, there were practically two standpoints as regards the status of infinitesimals. According to one of them, which seemed to be shared by Leibniz, an infinitely small quantity was thought of as an entity “smaller than any given or assignable magnitude.” Actual “indivisible” elements comprising numerical quantities and geometrical figures are the perceptions corresponding to this concept of the infinitely small. Leibniz did not doubt the existence of “simple substances incorporated into the structure of complex substances,” i.e., *monads*. “It is these monads that are the genuine atoms of nature or, to put it short, elements of things” [312, p. 413].

For the other founder of analysis, Newton, the concept of infinite smallness is primarily related to the idea of vanishing quantities [384, 408]. He viewed the indeterminate quantities “not as made up of indivisible particles but as described by a continuous motion” but rather “as increasing or decreasing by a perpetual motion, in their nascent or evanescent state.”

The celebrated “method of prime and ultimate ratios” reads in his classical treatise *Mathematical Principles of Natural Philosophy* (1687) as follows: “Quantities, and the ratios of quantities, which in any finite time converge continuously

to equality, and before the end of that time approach nearer to each other than by any given difference, become ultimately equal” [408, p. 101].

Propounding the ideas which are nowadays attributed to the theory of limits, Newton exhibited the insight, prudence, caution, and wisdom characteristic of any great scientist pondering over the concurrent views and opinions.

He wrote: “To institute an analysis after this manner in finite quantities and investigate the prime or ultimate ratios of these finite quantities when in their nascent state is consonant to the geometry of the ancients, and I was willing to show that in the method of fluxions there is no necessity of introducing infinitely small figures into geometry.

Yet the analysis may be performed in any kind of figure, whether finite or infinitely small, which are imagined similar to the evanescent figures, as likewise in the figures, which, by the method of indivisibles, used to be reckoned as infinitely small provided you proceed with due caution” [384, p. 169].

Leibniz’s views were as much pliable and in-depth dialectic. In his famous letter to Varignon of February 2, 1702 [408], stressing the idea that “it is unnecessary to make mathematical analysis depend on metaphysical controversies,” he pointed out the unity of the concurrent views of the objects of the new calculus:

“If any opponent tries to contradict this proposition, it follows from our calculus that the error will be less than any possible assignable error, since it is in our power to make this incomparably small magnitude small enough for this purpose, inasmuch as we can always take a magnitude as small as we wish. Perhaps this is what you mean, Sir, when you speak on the inexhaustible, and the rigorous demonstration of the infinitesimal calculus which we use undoubtedly is to be found here. ...

So it can also be said that infinites and infinitesimals are grounded in such a way that everything in geometry, and even in nature, takes place as if they were perfect realities. Witness not only our geometrical analysis of transcendental curves but also my law of continuity, in virtue of which it is permitted to consider rest as infinitely small motion (that is, as equivalent to a species of its own contradictory), and coincidence as infinitely small distance, equality as the last inequality, etc.”

Similar views were expressed by Leibniz in the following quotation whose end in italics is often cited in works on infinitesimal analysis in the wake of Robinson [421, pp. 260–261]:

“There is no need to take the infinite here rigorously, but only as when we say in optics that the rays of the sun come from a point infinitely distant, and thus are regarded as parallel. And when there are more degrees of infinity, or infinitely small, it is as the sphere of the earth is regarded as a point in respect to the distance of the sphere of the fixed stars, and a ball which we hold in the hand is also a point in comparison with the semidiameter of the sphere of the earth. And then the distance to the fixed stars is infinitely infinite or an infinity of infinities in relation

to the diameter of the ball. For in place of the infinite or the infinitely small we can take quantities as great or as small as is necessary in order that the error will be less than any given error. In this way *we only differ from the style of Archimedes in the expressions, which are more direct in our method and better adapted to the art of discovery.*" [311, p. 190].

1.2. L. Euler

The 18th century is rightfully called the age of Euler in the history of mathematical analysis (cf. [45]). Everyone looking through his textbooks [112] will be staggered by subtle technique and in-depth penetration into the essence of the subject. It is worth recalling that an outstanding Russian engineer and scientist Krylov went into raptures at the famous Euler formula $e^{i\pi} = -1$ viewing it as the quintessential symbol of integrity of all branches of mathematics. He noted in particular that "here 1 presents arithmetic; i , algebra; π , geometry; and e , analysis."

Euler demonstrated an open-minded approach, which might deserve the epithet "systemic" today, to studying mathematical problems: he applied the most sophisticated tools of his time. We must observe that part and parcel of his research was the effective and productive use of various infinitesimal concepts, first of all, infinitely large and infinitely small numbers. Euler thoroughly explained the methodological background of his technique in the form of the "calculus of zeros." It is a popular fixation to claim that nothing is perfect and to enjoy the imaginary failures and follies of the men of genius ("to look for sun-spots" in the words of a Russian saying). For many years Euler had been incriminated in the "incorrect" treatment of divergent series until his ideas were fully accepted at the turn of the 20th century. You may encounter such a phrase in the literature: "As to the problem of divergent series, Euler was sharing quite an up-to-date point of view." It would be more fair to topsy-turvy this phrase and say that the mathematicians of today have finally caught up with some of Euler's ideas. As will be shown in the sections to follow (see 2.2 and 2.3), the opinion that "we cannot admire the way Euler corroborates his analysis by introducing zeros of various orders" is as self-assured as the statement that "the giants of science, mainly Euler and Lagrange, have laid false foundations of analysis." It stands to reason to admit once and for ever that Euler was in full possession of analysis and completely aware what he had created.

1.3. G. Berkeley

The general ideas of analysis greatly affected the lineaments of the ideological outlook in the 18th century. The most vivid examples of the depth of penetration of the notions of infinitely large and infinitely small quantities into the cultural media of that time are in particular *Gulliver's Travels* by Jonathan Swift published

in 1726 (Lilliput and Brobdingnag) and the celebrated *Micromegas 1752* written by bright and venomous F. M. Arouer, i.e., Voltaire. Of interest is the fact that as an epigraph for his classical treatise [421], Robinson chose the beginning of the following speech of Micromegas:

“Now I see clearer than ever that nothing can be judged by its visible magnitude. Oh, my God, who granted reason to creatures of such tiny sizes! An infinitely small thing is equal to an infinitely large one when facing you; if living beings still smaller than those were possible, they could have reason exceeding the intellect of those magnificent creatures of yours which I can see in the sky, and one foot of which could cover the earth” [507, p. 154].

A serious and dramatic impact on the development of infinitesimal analysis was made in 1734 by Bishop Berkeley, a great cleric and theologian, who published the pamphlet *The Analyst, or a Discourse Addressed to an Infidel Mathematician, wherein it is examined whether the object, principles and inferences of the modern analysis are more deduced than religious mysteries and points of faith* [34]. By the way, this Infidel Mathematician was E. Halley, a brilliant astronomer and a young friend of Newton. The clerical spirit of this article by Berkeley is combined with aphoristic observations and killing precision of expression. The leitmotif of his criticism of analysis reads: “Error may bring forth truth, though it cannot bring forth science.”

Berkeley’s challenge was addressed to all natural sciences: “I have no controversy about your conclusions, but only about your logic and method. How do you demonstrate? What objects are you conversant with, and whether you conceive them clearly? What principles you proceed upon; how sound they may be; and how you apply them?” Berkeley’s invectives could not be left unanswered by the most progressive representatives of the scientific thought of the 18th century, the encyclopedists.

1.4. J. D’Alembert and L. Carnot

A turning point in the history of the basic notions of analysis is associated with the ideas and activities of D’Alembert, one of the initiators and leading authors of the immortal masterpiece of the thought of the Age of Enlightenment, the French *Encyclopedia or Explanatory Dictionary of Sciences, Arts, and Crafts*.

In the article “Differential” he wrote: “Newton never considered differential calculus to be some calculus of the infinitely small, but he rather viewed it as the method of prime and ultimate ratios” [408, p. 157]. D’Alembert was the first mathematician who declared that he had found the proof that the infinitely small “do exist neither in Nature nor in the assumptions of geometricians” (a quotation from his article “Infinitesimal” of 1759).

The D'Alembert standpoint in *Encyclopedia* contributed much to the formulation by the end of the 18th century of the understanding of an infinitesimal as a vanishing magnitude.

It seems worthy to recall in this respect the book by Carnot *Considerations on Metaphysics of the Infinitely Small* wherein he observed that “the notion of infinitesimal is less clear than that of limit implying nothing else but the difference between such a limit and the quantity whose ultimate value it provides.”

1.5. B. Bolzano, A. Cauchy, and K. Weierstrass

The 19th century was the time of building analysis over the theory of limits. Outstanding contribution to this process belongs to Bolzano, Cauchy, and Weierstrass whose achievements are mirrored in every traditional textbook on differential and integral calculus.

The new canon of rigor by Bolzano, the definition by Cauchy of an infinitely small quantity as a vanishing variable and, finally, the ε - δ -technique by Weierstrass are indispensable to the history of mathematical thought, becoming part and parcel of the modern culture.

It is worth observing (see [408]) that, giving a verbal definition of continuity, both Cauchy and Weierstrass chose practically the same words:

An infinitely small increment given to the variable produces
an infinitely small increment of the function itself.

CAUCHY

Infinitely small variations in the arguments correspond
to those of the function.

WEIERSTRASS

This coincidence witnesses the respectful desire of the noble authors to interrelate the new ideas with the views of their great predecessors.

Speculating about significance of the change of analytical views in the 19th century, we should always bear in mind the important observation by Severi [439, p. 113] who wrote: “This reconsideration, close to completion nowadays, has however no everlasting value most scientists believe in. Rigor itself is, in fact, a function of the amount of knowledge at each historical period, a function that corresponds to the manner in which science handles the truth.”

1.6. N. N. Luzin

The beginning of the 20th century in mathematics was marked by a growing distrust of the concept of infinitesimal. This tendency became prevailing as mathematics was reconstructed on the set-theoretic foundation whose proselytes gained the key strongholds in the 1930s.

In the first edition of the *Great Soviet Encyclopedia* in 1934, Luzin wrote: “As to a constant infinitely small quantity other than zero, the modern mathematical analysis, without discarding the formal possibility of defining the idea of a constant infinitesimal (for instance, as a corresponding segment in some non-Archimedean geometry), views this idea as absolutely fruitless since it turns out impossible to introduce such an infinitesimal into calculus” [335, pp. 293–294].

The publication of the textbook *Fundamentals of Infinitesimal Calculus* by Vygodskiĭ became a noticeable event in Russia at that time and gave rise to a serious and sharp criticism. Vygodskiĭ tried to preserve the concept of infinitesimal by appealing to history and paedeutics.

He wrote in particular: “If it were only the problem of creating some logical apparatus that could work by itself then, having eliminated infinitesimals from considerations and having driven differentials out of mathematics, one could celebrate a victory over the difficulties that have been impeded the way of mathematicians and philosophers during the last two centuries. Infinitesimal analysis originated however from practical needs, its relations with the natural sciences and technology (and, later, with social sciences) becoming increasingly strong and fruitful in the course of time. Complete elimination of infinitesimals would hinder these relations or even make them impossible” [515, p. 160].

Discussing this textbook by Vygodskiĭ, Luzin wrote in the 1940s: “This course, marked by internal integrity and lit by the great idea the author remains faithful to, falls beyond the framework of the style in which the modern mathematical analysis has been developed for 150 years and which is nearing its termination” [335, p. 398].

Luzin’s attitude to infinitesimals deserves special attention as apparent manifestation and convincing evidence of the background drama typical of the history of every profound idea that enchants and inspires the mankind. Luzin had a unique capability of penetration into the essence of the most intricate mathematical problems, and he might be said to possess a remarkable gift of foresight [308, 309, 337].

The concept of actual infinitesimals seemed to be extremely appealing to him psychologically, as he emphasized: “The idea about them has never been successfully driven out of my mind. There are obviously some deeply hidden reasons still unrevealed completely that make our minds inclined to looking at infinitesimals favorably” [335, p. 396].

In one of his letters to Vygodskiĭ which was written in 1934 he predicted that “infinitesimals will be fully rehabilitated from a perfectly scientific point of view as kind of ‘mathematical quanta.’”

In another of his publications, Luzin sorrowfully remarked: “When the mind starts acquaintance with analysis, i.e., during the mind’s spring season, it is the infinitesimals, which deserve to be called the ‘elements’ of quantity, that the mind begins with. However, surfeiting itself gradually with knowledge, theory, abstrac-

tion and fatigue, the mind gradually forgets its primary intentions, smiling at their ‘childishness.’ In short, when the mind is in its autumn season, it allows itself to become convinced of the unique sound foundation by means of limits” [504].

This limit conviction was energetically corroborated by Luzin in his textbook *Differential Calculus* wherein he particularly emphasized the idea that “to grasp the very *essence of the matter* correctly, the student should first of all made it clear that each infinitesimal is always a variable quantity *by its very definition*; therefore, no constant number, however tiny, is *ever* infinitely small. The student should beware of using comparisons or similes of such a kind for instance as ‘One centimeter is a magnitude infinitely small as compared with the diameter of the sun.’ This phrase is pretty incorrect. Both magnitudes, i.e., one centimeter and the diameter of the sun, are constant quantities and so they are *finite*, one much smaller than the other, though. Incidentally, one centimeter is not a small length at all as compared for instance with the ‘thickness of a hair,’ becoming a long distance for a moving microbe. In order to eliminate any risky comparisons and haphazard subjective similes, the student *must remember that neither constant magnitude is infinitesimal nor any number, however small these might be*. Therefore, it would be quite appropriate to abandon the term ‘*infinitesimal magnitude*’ in favor of the term ‘*infinitely vanishing variable*,’ as the latter expresses the *idea of variability* most vividly” [504, p. 61].

1.7. A. Robinson

The seventh (posthumous) edition of this textbook by Luzin was published in 1961 simultaneously with Robinson’s *Nonstandard Analysis* which laid a modern foundation for the calculus of infinitesimals. Robinson based his research on the local theorem by Mal’tsev, stressing its “fundamental importance for our theory” [421, p. 13] and giving explicit references to Mal’tsev’s article dated as far back as 1936. Robinson’s discovery elucidates the ideas of the founders of differential and integral calculus, witnessing the spiral evolution of mathematics.

Chapter 2

Naive Foundations of Infinitesimal Analysis

The most widely spread prejudice against infinitesimals resides in the opinion that the technique of infinitesimal analysis is extremely difficult to master. Moreover, it is usually emphasized that the nonstandard methods of analysis rest on rather sophisticated sections of set theory and mathematical logic. This circumstance is irrefutable but overrated, hampering in no way comprehension of infinitesimals.

The purpose of this chapter is to corroborate the above statement by presenting the methodology of infinitesimal analysis at the routine level of rigor which is offered by the modern system of mathematical education invoking the naive set-theoretic stance that stems from Cantor. Alongside with elucidating the basic concepts of nonstandard set theory and its principles of transfer, idealization, and standardization, we pay attention also to comparing the new views of the basic concepts of analysis with those of the reverent inventors of the past. We hope so to witness the continual evolution and immortality of the ideas of differential and integral calculus which infinitesimal analysis in a today's disguise shed new light upon.

2.1. The Concept of Set in Infinitesimal Analysis

In this section we will set forth a fragment of the foundations of infinitesimal analysis at the level of rigor close to the current practice of teaching calculus.

2.1.1. Contemporary courses in mathematical analysis rest usually on the concept of set.

2.1.2. EXAMPLES.

(1) L. Schwartz, *Analysis*:

“A set is a collection of objects.

Examples: the set of all alumni of a school;

the set of points on a plane;
 the set of nondegenerate surfaces
 of second-order in three-dimensional space;
 the set \mathbb{N} of positive integers;
 the set \mathbb{Z} of integers;
 the set \mathbb{Q} of rational numbers;
 the set \mathbb{R} of real numbers;
 the set \mathbb{C} of complex numbers" [436].

(2) V. A. Il'in, V. A. Sadovnichii, and Bl. Kh. Sendov, *Mathematical Analysis*:

"The concept of set was of importance when we have studied reals. We emphasize that we view a set as a basic concept indeterminate from the others.

In this section we will study sets of an arbitrary nature which are also called abstract sets. This implies that the objects comprising such a set which we call the elements of this set are not necessarily some real numbers. For instance, the arbitrary functions, letters of an alphabet, planar figures, etc. may serve as elements of an abstract set [189, p. 69].

(3) Yu. G. Reshetnyak, *A Course in Mathematical Analysis*:

"A set for us will be one of the basic mathematical concepts inexpressible in the other mathematical concepts. Uttering the word 'set,' we usually imply a collection of objects of an arbitrary nature which we will treat as a whole. Alongside this term, set, we will use its synonyms like totality, system, assembly, and so on. We may speak for instance about the set of solutions to an equation, about the collection of pictures in a museum, the totality of points of a circle and so on.

The objects, comprising a set, are the *elements* of this set.

We assume a set given if, granted whatever object, we can determine whether or not it is an element of the set in question" [413, p. 12].

(4) V. A. Zorich, *Mathematical Analysis*:

"The basic hypotheses of cantorion set theory (called 'naive' in common parlance) are as follows:

- 1° each set may consists of arbitrary distinct objects;
- 2° each set is uniquely determined by the collection of objects comprising it;
- 3° each property determines a set of objects enjoying this property.

If x is an object, P is a property, with $P(x)$ signifying that x enjoys P , then we let $\{x \mid P(x)\}$ stand for the class of objects possessing P .

The objects comprising a class or a set are the *elements* of the class or the set.

The words 'class,' 'family,' 'totality,' and 'collection' are viewed as synonyms of the noun 'set' within the naive set theory.

The following examples demonstrate the application of this terminology:

the set of letters 'a' in the word 'I';

the set of Adam's wives;
 the collection of ten digits;
 the family of Leguminosae;
 the set of grains of sand on the Earth;
 the totality of the points on a plane equidistant from two given points;
 a family of sets;
 the set of all sets.

The degree of certainty of definition varies greatly from set to set. This prompts us the conclusion that the concept of set is not as simple and harmless as it might seem. For example, the definition of the set of all sets leads in fact to an outright contradiction" [541, pp. 17–18].

2.1.3. Infinitesimal or nonstandard analysis belongs in mathematical analysis. Therefore, this discipline fully shares the above views of sets. Consequently, *infinitesimal analysis considers as sets those and only those collections that are admitted into the classical "standard" theory.*

It is worth observing that the last statement may be reformulated as follows: Infinitesimal analysis refuses to view as sets those and only those collections that are rejected by the present-day mathematics. At the same time, *infinitesimal analysis rests on refining views of sets.* To put it otherwise, infinitesimal analysis resides within the realm of nonstandard set theory.

2.1.4. The naive set theory starts with the celebrated definitions by Cantor: "A set is any many which can be thought of as one, that is every totality of definite elements which can be united to a whole through a law," and a set is "every collection into a whole M of definite and distinct objects m of our perception or our thought" [52, p. 173].

Such concepts are well known to be rather broad, this drawback bypassed by elaborating distinction between sets and nonsets. For instance, the term "class" is in common parlance for nominating "exceedingly huge" inappropriate collections, implying that a class is not necessarily a set. In other words, formalization of the concepts of the naive set theory rests on clarifying and elaborating the procedures that introduce sets into the practice of mathematics. All sets we admit into mathematics enjoy the same rights. This in no way implies that all sets are equal or bear no distinctions but means simply that all sets are akin, maintaining the same status of an ordinary member of the "universe of sets."

2.1.5. The cornerstone of nonstandard set theory is the following perfectly transparent underlying principle: *The sets differ from each other: every set is either standard or nonstandard.* In a way, it would be more correct to speak of the theory of sets, standard and nonstandard, rather than nonstandard set theory.

The phrase "A is a standard set" conveys the intuitive idea that A admits

description in accurate and unambiguous terms. We may even say that such an A becomes an “artifact” of the cognition activities of human beings. The concept of standardness draws a borderline between the objects resulting from explicit mathematical constructions, primarily, by the theorems of unique existence, which are *standard sets*, and the objects originating in research in implicit, indirect manner, which are *nonstandard sets*.

Such objects as the numbers π , e , and $\sin 81$ are determined uniquely in much the same way as the set of natural numbers \mathbb{N} and the set \mathbb{R} of real numbers also called the *reals*. These are standard objects. On the other hand, an arbitrary “abstract” real number arises implicitly on assuming the set-theoretic stance of mathematics. Such an “abstract” real number is by definition just a member of the standard set \mathbb{R} , the reals. This is a routine method of introducing objects into mathematical practice: a *vector* is an element of a vector space; a *filter* is a set of some subsets of a given set which enjoy a few specific properties; etc.

Therefore, there are standard reals and nonstandard reals; there are standard vectors and nonstandard vectors; there are standard filters and nonstandard filters; etc. In general, there are standard sets and nonstandard sets.

By way of example, let us consider the set of grains of sand on the Earth.

Archimedes wrote in his classical treatise *Psammiths, the Sand-Reckoner*, that: “. . . of the numbers named by me and given in the work which I sent to Zeuxippes some exceed not only the number of the grains of sand equal in magnitude to the Earth filled up in the way described but also that of a mass equal in magnitude to the *universe*” [2, p. 358].

Therefore, the number of the grains of sand on the Earth presents a particular natural number. However, nobody can either determine or nominate this number precisely, since it is absolutely impossible to implement any sequential count of all grains of sand. Hence, the number of the grains of sand on the Earth is expressed by an “*inassignable*,” “*indeterminate*”—nonstandard—natural number and so the set of the grains of sand is nonstandard.

It goes without saying that the above *views of distinction between standard and nonstandard sets are auxiliaries* for mastering the rules of handling these sets in practice. We encounter a complete analogy with the situation in geometry where the intuitive visualizations of spatial forms help in elaborating the skills of using the axioms of geometry which, in the long run, are the only source of the rigorous definitions of points, straight lines, planes, and other geometrical objects. Following Alexandrov, we must observe that “axioms themselves need substantiation since they only summarize the available data, while starting the logical construction of a theory” [7, p. 51]. Therefore, we are impelled to precede the formal introduction of the axioms of nonstandard set theory by discussing them qualitatively.

As we already know, any nonstandard set theory begins with the primary obser-

vation that we distinguish two types of sets: standard and nonstandard. Moreover, we accept another three postulates or, to be more exact, versions of the principles of infinitesimal analysis.

2.1.6. Transfer Principle. *Each mathematical proposition claiming the existence of some set simultaneously determines a standard set.*

In other words, the unique existence theorems of mathematics provide the direct and explicit definitions of new mathematical entities. An equivalent reformulation of this principle, elucidating the reasons behind its official title, reads: *In order to validate some claim about all sets, it suffices to demonstrate the claim in the case of standard sets.* Intuitive substantiation of the transfer principle lies with the evident fact that, making every judgement about arbitrary sets, we actually deal only with the already available sets we have defined uniquely, i.e., with standard sets.

Pondering over the meaning of the transfer principle, we see that it contains two aspects of the views of standard entities. The first proclaims that *new standard objects result from those already available by descriptions similar to the theorems of unique existence*, which postulates the possibility of defining standard objects by *recursion*. This circumstance may be rephrased as the conclusion that each nonempty standard set contains a standard element and each entity is standard that results from standard elements by some unique construction or definition.

The second aspect of views of standardness, as expressed by the transfer principle is interwoven with the first, and means *representativeness of the standard universe*, i.e. the class of standard sets is sufficiently large for reflecting all features of the universe of sets. In other words, this postulates the possibility of studying arbitrary sets by *induction* starting from actually available standard entities, i.e., cognizability of ideal constructions of set theory.

2.1.7. Idealization Principle. *Each infinite set contains a nonstandard element.*

This postulate conforms perfectly with the most general views of infinity. The idealization principle will appear often in stronger forms reflecting the inexhaustible variety of ideal objects. For instance, one of the popular nonstandard axioms reads: *All standard sets are members of some finite set.* The number of elements of such a “universal” set is huge and, which is most important, “inassignable,” “inexpressible,” and “unrealizable.” In other words, the size of a universal set is nonstandard and so it is no surprise that every universal set is nonstandard either.

It is worth observing that care must be exercised in handling the above postulates as on the other occasions by the way. Indeed, every standard set is uniquely determined from its standard elements in the *standard environment*—in the community of its next of kin, other standard sets. However, an infinite standard set

never reduces to its standard elements since it always contains some nonstandard element. There are many nonstandard sets containing each standard element of the original set and having no other standard elements.

Another circumstance is very important to mention: *the term “proposition” in the transfer principle deserves extreme caution* in much the same way as in the conventional set theory although. Transfer is fully legitimate inasmuch as we apply it to the common mathematical proposition not involving the property of a set to be or not to be standard which we have introduced at a semantic level in the capacity of another indeterminate basic concept of nonstandard set theory. Indeed, were it otherwise, we would apply transfer to the claim that every set is standard, arguing as follows: Since every standard set is standard, every set is standard by transfer. This would yield a contradiction with idealization. Therefore, transfer is not applicable to the claim that every set is standard and so the standardness predicate is a property inexpressible in the naive set theory.

2.1.8. Standardization Principle. *To each standard set and each property there corresponds a new standard set comprising precisely those standard elements of the original set that possess the property under study.*

In symbols, let A be a standard set, and let P be an arbitrary property which may involve the standardness predicate in its formulation. The standardization principle claims that there is a standard set which is usually denoted by $^*\{x \in A : P(x)\}$ and maintains the relation

$$y \in ^*\{x \in A : P(x)\} \leftrightarrow y \in \{x \in A : P(x)\}$$

for all standard y . The set $^*\{x \in A : P(x)\}$ is often referred to as the *standardization* of A or even briefer. We write *A , omitting the parameters that participate in the definition of the standardization of A . The intuitive idea behind the standardization principle reflects the experience showing that if some explicit descriptions of mathematical objects are available then we may use any definite rules for assigning new entities of further mathematical research. *Standardization extends the conventional comprehension principle of set formation which allows us to deal with a new subset of a given set by collecting the elements with some prescribed property.*

Thinking over the standardization principle, it is worth noting that nothing is claimed about the nonstandard elements of the result of standardization. This is not by chance. The point is that the two possibilities are open: a nonstandard element can enjoy the property we use in standardization or it may fail to possess this property.

The standardization principle must be used with the same care as elsewhere: Attempts at standardizing some “universal” set that contains all standard sets would result in a blatant contradiction.

2.1.9. The above postulates give grounds for axiomatic presentations of non-standard set theory. We will discuss these in more detail in Chapter 3. Meanwhile, we may proceed along the lines of the Zorich textbook and remark: “As a whole, each of the available axiomatizations is such that, on the one hand, it eliminates the known contradictions of the naive theory and, on the other hand, it ensures freedom of operation with concrete sets residing in various sections of mathematics and, before all, in mathematical analysis understood in a broad sense of the word” [541, pp. 18–19].

2.2. Preliminaries on Standard and Nonstandard Reals

We now start an acquaintanceship with the surprising features of the classical real axis which are disclosed by the principles of infinitesimal analysis.

2.2.1. Given a set A , we will write $a \in {}^\circ A$ to abbreviate the expression “ a is a standard member of A .”

2.2.2. *The following hold:*

- (1) *If A is a set satisfying $1 \in A$ and such that $n \in {}^\circ \mathbb{N}$ implies $n \in A \rightarrow n + 1 \in A$ then A contains all standard natural numbers: ${}^\circ \mathbb{N} \subset A$ (= the induction principle on standard naturals);*
- (2) *Every finite set (i.e., a set admitting no injective mappings onto any of its proper subsets), consisting of standard elements, is standard itself;*
- (3) *Every finite standard set has only standard elements;*
- (4) *If a set contains only standard members then it is finite;*
- (5) *The totality ${}^\circ A$ is not a set for every infinite (= not finite) set A .*

◁ (1): By standardization, it is possible to consider the following standard subset of \mathbb{N} , the set of natural numbers:

$$B := {}^* \{n \in \mathbb{N} : n \notin A\}.$$

Assuming that $B \neq \emptyset$, observe that B has the least element m which is standard by transfer. By hypothesis, $m \neq 1$ since $1 \in A$. Moreover, $m \notin A$ and so $m - 1 \notin A$. By transfer, $m - 1 \in {}^\circ \mathbb{N}$, i.e., $m - 1 \in B$. It follows that $m - 1 \geq m$, which is a contradiction. Hence, $B = \emptyset$, i.e., $(\forall n \in {}^\circ \mathbb{N})(n \in A)$, implying the inclusion ${}^\circ \mathbb{N} \subset A$.

(2): This is immediate by transfer.

(3): Each standard singleton contains a unique and, hence, standard element. The *size*, i.e. the number of elements, of a finite standard set A is also standard by transfer. Moreover, $A = (A - \{a\}) \cup \{a\}$ for all $a \in A$. The number of elements of $A - \{a\}$ is also standard, and we are done by the induction principle (1).

(4): This is straightforward by idealization.

(5): Assume that ${}^\circ A$ is a set. The set ${}^\circ A$ is finite by (4) and standard by (2). By transfer, $A = A^\circ$ and so A is finite, which is a contradiction. \triangleright

2.2.3. *A natural number N is nonstandard (i.e., inassignable) if and only if N is greater than every standard natural number. In symbols,*

$$N \in \mathbb{N} - {}^\circ\mathbb{N} \leftrightarrow (\forall n \in {}^\circ\mathbb{N})(N > n).$$

\triangleleft It suffices to note that if $n \in \mathbb{N}$ is standard and $N \in \mathbb{N}$ is less than or equal to n then N is standard for instance by 2.2.2, i.e., $N \in {}^\circ\mathbb{N}$. \triangleright

2.2.4. In view of 2.2.3, a nonstandard natural number is called *unlimited*, *illimited*, *actually infinite*, or even *infinite*, which abuses the language.

It is a widely-spread opinion that “Euler claimed rather light-mindedly that $1/0$ is infinite, although he never found it worthwhile to define what infinity is; he just invented the notation ∞ .” This opinion is clearly false since Euler pointed out explicitly in [110, p. 89] that “an infinite number and a number greater than any assignable number are synonyms.”

The fact that N is an unlimited number is expressed symbolically as $N \approx \infty$ or, in more detail, $N \approx +\infty$.

We must mention that the epithet “infinite” for an unlimited number N leads to confusion. Indeed, if we strictly pursue the set-theoretic stance then we view N primarily as a set and this set N is clearly finite in the rigorous set-theoretic sense (cf. 2.2.2(2)). The phrase “ N is an infinite number” suggests misleadingly that N is an infinite set. In actuality, N is a finite set whose size is nonstandard. Only this meaning is implied in the concept of an infinitely large natural number N within the set-theoretic credo of contemporary mathematics.

2.2.5. *The following hold:*

- (1) $(N \approx +\infty \text{ and } M \approx +\infty) \rightarrow (N + M \approx +\infty \text{ and } NM \approx +\infty)$;
- (2) $(N \approx +\infty \text{ and } n \in {}^\circ\mathbb{N}) \rightarrow (N + n \approx +\infty, N - n \approx +\infty, \text{ and } nN \approx +\infty)$;
- (3) $N \approx +\infty \leftrightarrow N^n \approx +\infty$ for all $n \in {}^\circ\mathbb{N}$;
- (4) *Each composite unlimited number has unlimited divisors;*
- (5) $(N \approx +\infty \text{ and } M \geq N) \rightarrow M \approx +\infty$;
- (6) “If $\frac{1}{0}$ denotes some infinitely large number then, since $\frac{2}{0}$ is undoubtedly the doubled $\frac{1}{0}$, it is clear that every number, even infinitely large, can be made still two or several times greater” (Euler, [109, p. 620]);
- (7) *Let t be a positive real. The integral part of t is unlimited if and only if so is t , i.e., t is greater than every standard real;*

(8) Let $\psi : \mathbb{N} \rightarrow \mathbb{N}$ be a strictly increasing standard function. Then

$$N \approx +\infty \leftrightarrow \psi(N) \approx +\infty$$

for $N \in \mathbb{N}$.

◁ We will demonstrate only (7) and (8), since the other claims are easier to check.

(7): If the integral part s of a real t is unlimited and, nevertheless, $(\exists r \in {}^\circ\mathbb{R}) (t \leq r)$ then $t \leq n$ for some $n \in {}^\circ\mathbb{N}$. Hence, $n + 2 \leq s \leq t \leq n$, which is a contradiction. Therefore, $t \approx +\infty$.

Conversely, if $t \approx +\infty$ then $s + 1 \geq t$, where s is the integral part of t . Hence, $s + 1 \approx +\infty$, which yields $s \approx +\infty$ by 2.2.5(2).

(8): ←: Assume first that $\psi(N) \approx +\infty$ and $n \in {}^\circ\mathbb{N}$. Then $\psi(n)$ is standard by transfer, i.e., $\psi(n) \in {}^\circ\mathbb{N}$. Therefore, $\psi(N) > \psi(n)$. Since ψ increases strictly, it follows that $N > n$, i.e., $N \approx +\infty$.

→: Assume now that $N \approx +\infty$. Then $N > n$ for all $n \in {}^\circ\mathbb{N}$. Hence, $\psi(N) > \psi(n) \geq n$. Thus, $\psi(N) \approx +\infty$. ▷

2.2.6. Let $\overline{\mathbb{R}}$ stand for the *extended reals*, i.e., $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ with $+\infty$ and $-\infty$ the greatest and least elements appended to \mathbb{R} formally. It is convenient to call $\infty := \{+\infty, -\infty\}$ the (symbolic) *potential infinity* and to speak of $+\infty$ or $-\infty$ as about the positive or negative (symbolic) infinity.

A real $t \in \mathbb{R}$ is *limited* provided that there is a standard number $n \in {}^\circ\mathbb{N}$ satisfying $|t| \leq n$. We write $t \in \text{ltd}(\mathbb{R})$ or $t \in \approx\mathbb{R}$ whenever t is a limited element of \mathbb{R} . A member of \mathbb{R} that fails to be limited is *unlimited* or *actually large*. We also write $t \approx +\infty$ for $t \notin \approx\mathbb{R}$ and $t > 0$. The records $t \approx -\infty$ and $t \approx \infty$ are understood by analogy. It is a common practice to agree that $t \approx +\infty \leftrightarrow t \in \mu(+\infty)$, using the expressions like “ t lies in the *monad* of the point at infinity or the monad of the plus-infinity.”

An element t in \mathbb{R} is *infinitesimal* or, amply, *infinitely small* provided that $|t| \leq 1/n$ for all $n \in {}^\circ\mathbb{N}$. In this event we write $t \approx 0$ or $t \in \mu(\mathbb{R})$ and say that t belongs to the *monad* of zero. (The symbol $\mu(\mathbb{R})$ is common alongside the notation $\mu(0)$ which signifies a close relationship with the unique separated vector topology on \mathbb{R} .) Infinitesimals are also referred to as *actual infinitely small quantities*; the unsuccessful term “differentials” have little room.

If $x \leq y$ and the difference between x and y fails to be infinitesimal then we write $x \ll y$. Since $t \in \approx\mathbb{R} \leftrightarrow (\forall N \approx +\infty)(|t| \ll N)$; therefore, we let the record $|t| \ll +\infty$ symbolize the fact that $t \in \mathbb{R}$ is a limited real.

2.2.7. The term *monad* (Μόνναζ) dates from antiquity. It is traditionally translated as one or unit with no sufficient ground for that. By Definition 1 of

Book VII of Euclid's *Elements*, a monad "is (that) by virtue of which each of the things that exist is called one" [108, p. 9].

We now exhibit some qualitative observations about the structure of a monad as expressed by Sextus Empiricus:

"Pythagoras said that the origin of the things that exist is a monad by virtue of which each of the things that exist is called one" [440, p. 361];

"A point is structured as a monad; indeed, a monad is a certain origin of numbers and likewise a point is a certain origin of lines" [440, p. 364];

"A whole as such is indivisible and a monad, since it is a monad, is not divisible. Or, if it splits into many pieces it becomes a union of many monads rather than a [simple] monad" [440, p. 367].

We will study the status and structure of monads in great detail later. We now start with considering the basic properties of the collection of infinitesimals or, which is the same, the monad comprising all infinitely small reals.

2.2.8. *The following are valid:*

(1) $(s \approx 0 \text{ and } t \approx 0) \rightarrow s + t \approx 0;$

(2) $(t \approx 0 \text{ and } s \in \approx\mathbb{R}) \rightarrow st \approx 0;$

(3) $z \approx 0 \leftrightarrow 1/z \approx \infty$ (for $z \neq 0$):

"If z becomes a magnitude less than every assignable magnitude whatever, i.e., infinitely small, then the value of the fraction $\frac{1}{z}$ must become greater than every assignable magnitude, i.e., an infinitely large magnitude" (Euler [110, p. 93]).

(4) *If $t \approx 0$ and t is standard then $t = 0$.*

\triangleleft (1): Take $n \in {}^\circ\mathbb{N}$. Clearly, $|s| \leq (2n)^{-1}$ and $|t| \leq (2n)^{-1}$. Hence, $|s + t| \leq |s| + |t| \leq (2n)^{-1} + (2n)^{-1} = n^{-1}$, and so $s + t$ is infinitesimal.

(2): Assume that $s \in \approx\mathbb{R}$ and $s \neq 0$ (otherwise there is nothing left to prove). Take $n \in {}^\circ\mathbb{N}$. By hypothesis, $|s| \leq m$ for some $m \in {}^\circ\mathbb{N}$. Therefore, $|t| \leq (nm)^{-1}$. It follows that $|st| \leq |s| |t| \leq m(nm)^{-1} = n^{-1}$, i.e., $st \approx 0$.

(3): Let z be a limited nonzero real; i.e., $0 < |z| \leq |n|$, where $n \in {}^\circ\mathbb{N}$. Obviously, $|1/z| \geq 1/n$, i.e., $1/z$ is not infinitesimal. Conversely, if $z \approx \infty$ then $|z| \geq n$ for every limited n , which implies $z^{-1} \approx 0$.

(4): Note that $|t| \leq 2^{-1}|t|$ provided that t is standard, which is impossible for $|t| > 0$. Hence, $t = 0$. \triangleright

2.2.9. *The monad $\mu(\mathbb{R})$ is not a set.*

\triangleleft Assume the contrary. Then $\mu(\mathbb{R})$ is a subset of \mathbb{R} . Moreover, $t \geq \mu(\mathbb{R})$ for all $t > 0$, $t \in {}^\circ\mathbb{R}$. Hence, $t \geq s := \sup \mu(\mathbb{R})$. Obviously, s is a nonzero infinitesimal. However, $2s \geq s \rightarrow s = 0$, which contradicts the existence of nonzero infinitesimals. \triangleright

2.2.10. When we deal with reals, it is convenient to distinguish various cases of their interlocation.

Given $s, t, r \in \mathbb{R}$, we write $s = {}_r t$ or $s \approx t(\text{mod } r)$ provided that $(s - t)/r \approx 0$ (here $r \neq 0$). In this event we say that s and t are *r-close* or *infinitely close modulo r*. In case $r = 1$, we simply write $s \approx t$ and call s and t *infinitely close*.

The founders of infinitesimal analysis often made no distinction between the reals infinitely close to a given real and this real itself. Euler expressed this as follows: “An infinitely small quantity is exactly equal to zero” (or, in another translation, “an infinitely small quantity actually will be = 0”) [110, p. 92]. This explains the record $x = y$ for $x \in \mathbb{R}$ and $y \in {}^\circ\mathbb{R}$ satisfying $x \approx y$. In this respect Leibniz remarked: “I think that those things are equal not only whose difference is absolutely nothing but also whose difference is incomparably small” [311, p. 188], emphasizing that “the error is inassignable and cannot be found by means of whatever construction” [534, p. 195].

Given $s, t \in \mathbb{R}$, we write $s = O(t)$ for $s/t \in \approx\mathbb{R}$. In case $s = O(t)$ and $t = O(s)$ we then say that s and t have the *same order*. If $s/t \approx 0$ then we write $s = o(t)$ and say that s is of *higher order* than t . Finally, if $s - t = o(t)$ and $s - t = o(s)$ then we call s and t *equivalent* and write $s \sim t$.

Presenting his views on higher order infinitesimals, Leibniz wrote: “I would like to add one more remark in order to prevent all arguments against the reality of differences of any orders, and, namely, that they can always be expressed in proportional finite straight lines... I have already explained how to express first-order differences with conventional finite straight lines when first presenting the elements of this calculus in ‘Acta’ in October 1684” (see [534, pp. 188–190], cf. 1.1).

2.2.11. Given $s, t \in \mathbb{R}$, put

$$s \in O := O(t) \leftrightarrow s = O(t); \quad s \in o := o(t) \leftrightarrow s = o(t).$$

The Landau rules hold:

$$\begin{aligned} O + O &\subset O; & O + o &\subset O; & o + o &\subset o; \\ Oo &\subset o; & OO &\subset O; & oo &\subset o. \end{aligned}$$

◁ For definiteness, demonstrate the relation $O + o \subset O$. To this end, assume that $s := O(t)$ and $r := o(t)$. Then $s/t \in \approx\mathbb{R}$ and $r/t \approx 0$. Hence, $(s + r)/t \in \approx\mathbb{R}$, i.e., $s + r = O(t)$. ▷

2.2.12. Given $s, t \in \mathbb{R}$, the following are equivalent:

- (1) $s \sim t$;
- (2) $s - t = o(t)$ or $t - s = o(s)$;
- (3) $s/t \approx 1$ or $t/s \approx 1$;
- (4) $s/t \approx 1$ and $t/s \approx 1$.

◁ It is clear that (1) \rightarrow (2). If, for instance, $t - s = o(s)$ then $(t - s)/s \approx 0$; i.e., $t/s - 1 \approx 0$. Hence, $1 - \varepsilon \leq t/s \leq 1 + \varepsilon$ for $\varepsilon > 0$ and $\varepsilon \in {}^\circ\mathbb{R}$. Therefore, $(1 - \varepsilon)^{-1} \geq s/t \geq (1 + \varepsilon)^{-1}$ and $\varepsilon/(1 - \varepsilon) \geq s/t - 1 \geq \varepsilon/(1 + \varepsilon)$, i.e., $s/t \approx 1$. We have thus demonstrated that (2) \rightarrow (3) \rightarrow (4), while the implication (4) \rightarrow (1) is obvious. ▷

2.2.13. Given $N \in \mathbb{N}$, assume that $\alpha_k, \beta_k \in o(1)$ are infinitesimals satisfying $\alpha_k \sim \beta_k$ for $k := 1, \dots, N$. Then

- (1) $\sum_{k=1}^N \alpha_k \sim \sum_{k=1}^N \beta_k$ for $\alpha_k, \beta_k \geq 0$;
 (2) If $\sum_{k=1}^N |\alpha_k| = O(1)$ then

$$\sum_{k=1}^N \alpha_k \approx \sum_{k=1}^N \beta_k.$$

◁ To prove, note that 2.2.12 yields $-\varepsilon\alpha_k + \alpha_k \leq \beta_k \leq \alpha_k + \varepsilon\alpha_k$ for every standard $\varepsilon > 0$, which implies (1). Moreover, if $t := \sum_{k=1}^N |\alpha_k| \in {}^\circ\mathbb{R}$ then

$$\left| \sum_{k=1}^N (\alpha_k - \beta_k) \right| \leq \sum_{k=1}^N |\alpha_k - \beta_k| \leq \frac{\varepsilon}{n} \sum_{k=1}^N |\alpha_k| \leq \varepsilon$$

provided that $n \in {}^\circ\mathbb{N}$ satisfies $|t| \leq n$. ▷

2.2.14. There is a natural number N such that for each standard real t in \mathbb{R} the product Nt is infinitely close to some natural number.

◁ Consider some finite subset $\{x_1, \dots, x_n\}$ of \mathbb{R} that contains all standard reals, and take some positive infinitesimal $\varepsilon > 0$, $\varepsilon \approx 0$. The Dirichlet principle of number theory reads: To all $\varepsilon > 0$ and $x_1, \dots, x_n \in \mathbb{R}$ there is an integer $N \in \mathbb{N}$ such that each of the products Nx_1, \dots, Nx_n differs from some integers by at most ε . We are done on to apply this theorem with the chosen parameters. ▷

2.2.15. It is worth observing that the infinite proximity (as well as equivalence) of reals is not a subset of the cartesian square $\mathbb{R} \times \mathbb{R}$. Were it otherwise, the image of zero under this relation, which is the monad $\mu(\mathbb{R})$, would be a set which is impossible by 2.2.9. Note also that the monad $\mu(\mathbb{R})$ is *indivisible* in the following implicit sense: $n^{-1}\mu(\mathbb{R}) = \mu(\mathbb{R})$ for every standard n .

Pondering over the role of the monad $\mu(\mathbb{R})$ in the construction of the system of integers, it stands to reason to reconsider Definition 2 of Book VII of Euclid's *Elements* which ends as follows: "... and a number is a magnitude composed of monads" [108, p. 9].

Inspecting the "nonstandard" extended reals $\overline{\mathbb{R}}$, we see this set and, which is most nontrivial, its limited part $\approx\mathbb{R}$ to be the collection of monads about standard points. A more rigorous formulation of this statement rests on the following fundamental fact whose proof leans essentially on the standardization principle.

2.2.16. To each limited real t there corresponds a unique standard real infinitely close to t .

◁ Given $t \in {}^{\approx}\mathbb{R}$, use standardization and find the standard set $A := \{a \in \mathbb{R} : a \leq t\}$. Obviously, $A \neq \emptyset$ and $A \leq n$ whenever $n \in {}^{\circ}\mathbb{N}$ satisfies $-n \leq t \leq n$. Indeed, $a \leq t \leq n$ for every standard $a \in A$ and so $A \leq n$ by transfer. Since \mathbb{R} is complete; therefore, A has a unique least upper bound, $s := \sup A \in \mathbb{R}$. Obviously, s is standard by transfer. We will demonstrate that $s \approx t$. If not, we would have $|s - t| > \varepsilon$ for some standard $\varepsilon > 0$. If $s > t$ then $s > t + \varepsilon$ and so $s \geq a + \varepsilon$ for every standard a in A . Hence $s \geq s + \varepsilon$, which is impossible. The remaining possibility $s < t$ also leads to a contradiction since then we would find that $t > s + \varepsilon$ and, again, $s \geq s + \varepsilon$. ▷

2.2.17. The standard real infinitely close to a limited real $t \in {}^{\approx}\mathbb{R}$ is the *standard part* or *shadow* of t . The standard part of t is denoted by $\text{st}(t)$ or ${}^{\circ}t$.

We also agree for the sake of uniformity and convenience that ${}^{\circ}t = \text{st}(t) = +\infty$ whenever $t \approx +\infty$ and ${}^{\circ}t = \text{st}(t) = -\infty$ whenever $t \approx -\infty$ (on assuming as usual that $+\infty \approx +\infty$ and $-\infty \approx -\infty$). Therefore, to each (standard) $t \in \overline{\mathbb{R}}$ we put into correspondence the monad $\mu(t)$ of t comprising the elements s of \mathbb{R} for which $s \approx t$.

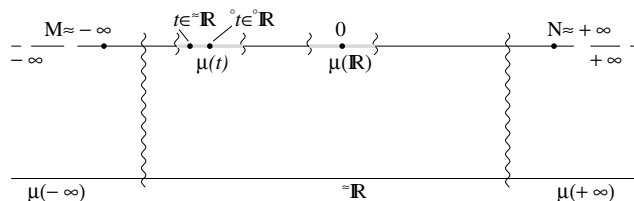


Fig. 2

We may summarize the above as follows: Infinitesimal analysis imagines the extended reals as shown in Fig. 2. Distinguishing a standard number ${}^{\circ}t$ on the axis \mathbb{R} , we draw a big dot, a blob, to symbolize the monad $\mu({}^{\circ}t)$ which is the “indivisible and explicit” image of ${}^{\circ}t$. Observing the region about t with a strong microscope, we will see in the eyepiece a blurred and dispersed cloud with unclear frontiers which is a visualization of $\mu(t)$.

Under greater magnification, the portion of the “point-monad” we are looking at will enlarge, revealing extra details whereas disappearing partially from sight. However, we are still inspecting the same standard real which you might prefer to percept as described by this process of “studying the microstructure of a physical straight line.”

2.2.18. The following hold:

(1) If $s \in \mathbb{R}$ and $t \in {}^{\approx}\mathbb{R}$ then

$$\text{st}(s + t) = \text{st}(s) + \text{st}(t); \quad \text{st}(st) = \text{st}(s)\text{st}(t);$$

- (2) If $s, t \in \mathbb{R}$ and $s \leq t$ then ${}^\circ s \leq {}^\circ t$;
 (3) If $s, t \in \overline{\mathbb{R}}$ then

$$(\exists t' \approx t)(t' \geq s) \leftrightarrow {}^\circ s \leq {}^\circ t \leftrightarrow (\forall \varepsilon > 0, \varepsilon \in {}^\circ \mathbb{R})(s \leq {}^\circ t + \varepsilon);$$

$$(\forall t' \approx t)(t' \geq s) \leftrightarrow {}^\circ s < t \quad ({}^\circ t \in {}^\circ \mathbb{R});$$

- (4) The standard part operation over the reals is not a set (hence, nor a function).

◁ (1): Prove by way of illustration that the standard part operation is multiplicative. Clearly, $s \approx \text{st}(s) \rightarrow ts \approx t\text{st}(s)$. Moreover, $t \approx \text{st}(t) \rightarrow \text{st}(s)t \approx \text{st}(t)\text{st}(s)$ which yields, $st \approx \text{st}(s)\text{st}(t)$. We are done on recalling that the product of standard reals is standard too.

(2): Assume that $s < t$ (otherwise everything is obvious). If $s \approx t$ then $\text{st}(s) = \text{st}(t)$. Were this otherwise, the monads $\mu(s)$ and $\mu(t)$ would be disjoint. Hence, ${}^\circ s < {}^\circ t$.

(3): In the initial equivalence, the implication to the right is obvious, while the reverse ensues from the fact that $s \leq {}^\circ t + s - {}^\circ s$ for $s \leq {}^\circ t$. Moreover, $s < t + \varepsilon \rightarrow \text{st}(s) \leq \text{st}(t) + \text{st}(\varepsilon) = {}^\circ t + \varepsilon$ for every $\varepsilon > 0, \varepsilon \in {}^\circ \mathbb{R}$. This implies by transfer that ${}^\circ s \leq {}^\circ t + \varepsilon$ for all positive ε . Hence, ${}^\circ s \leq {}^\circ t$.

Assume conversely that ${}^\circ s < {}^\circ t$. Since the monads $\mu({}^\circ s)$ and $\mu({}^\circ t)$ are disjoint; therefore, $s < {}^\circ t + \varepsilon$ for all $\varepsilon > 0, \varepsilon \in {}^\circ \mathbb{R}$.

To check the arrow to the right in the lower equivalence, note that s does not belong to the monad $\mu(t)$ of t . Hence, the whole of the monad of s lies to the left from the monad of t ; in symbols, $\mu(s) < \mu(t)$. Therefore, ${}^\circ s < {}^\circ t$. To prove the remaining implication, observe finally that $\mu(t) > {}^\circ s$ or $t \in \approx \mathbb{R}$ whenever ${}^\circ s = -\infty$. If ${}^\circ s \in {}^\circ \mathbb{R}$ then $\mu({}^\circ s) < {}^\circ t$. Hence, $t' \geq s$ whenever $t' \approx t$.

(4): If the law $t \mapsto \text{st}(t)$ were a set then the monad $\mu(\mathbb{R})$ would also be a set (since $t \in \mu(\mathbb{R}) \leftrightarrow {}^\circ t = 0$). It suffices to appeal to 2.2.9. ▷

2.3. Basics of Calculus on the Real Axis

We now discuss the fundamental notions of differential and integral calculus for functions in a single real variable.

2.3.1. Theorem. If (a_n) is a standard sequence and $a \in \overline{\mathbb{R}}$ is a standard real then

- (1) a is a partial limit of (a_n) if and only if $a = {}^\circ a_N$ for some unlimited natural N ;
 (2) a is the limit of (a_n) if and only if a_N is infinitely close to a for all unlimited naturals N ; in symbols,

$$a = \lim a_n \leftrightarrow (\forall N \approx +\infty)(a_N \approx a).$$

◁ These claims are demonstrated similarly. Therefore, we prove one of them, for instance, (2).

Assume for the sake of definiteness that $a_n \rightarrow a$ and $a \in \mathbb{R}$ (the cases $a = +\infty$ and $a = -\infty$ are settled by the same scheme). To each $\varepsilon > 0$ there is some $n \in \mathbb{N}$ such that $|a_N - a| \leq \varepsilon$ whenever $N \in \mathbb{N}$ and $n \leq N$. By transfer, to each standard $\varepsilon > 0$ there is some standard n with the same property. Every unlimited N is greater than n and so $|a_N - a| \leq \varepsilon$. Since ε is arbitrary, it follows that $a_N \approx a$.

Assume conversely that ${}^\circ a_N = a$ for all $N \approx +\infty$ and, for the sake of definiteness and diversity, $a := -\infty$. Take an arbitrary standard number $n \in {}^\circ\mathbb{N}$. If $N \geq M$ with M unlimited then $a_N \leq -n$. Given an arbitrary standard n , we have thus proved “something” (namely, “something” = $(\exists M)(\forall N \geq M)(a_N \leq -n)$). By transfer, this “something” holds for all $n \in \mathbb{N}$, which means clearly that $a_n \rightarrow -\infty$. ▷

2.3.2. We proudly note the merits of the above tests. We have demonstrated that the partial limits of a standard sequence are exactly those “assignable” reals that correspond to unlimited indices. In other words, each partial limit of (a_n) is the “observable” value of some remote entry of (a_n) . The tests of Theorem 2.3.1 in the standard environment have a clear intuitive meaning in contrast to the conventional definitions of partial limit as the limit of some subsequence of (a_n) or as such a point of the real axis whose every neighborhood intersects with every “tail” of (a_n) .

It is illuminative to look at the explanation of the concept of a partial limit of a [generalized] sequence with which Luzin furnished the formulation of the usual definition (see [334, pp. 98–99]): “The reader will undoubtedly find this definition cumbersome and abstract at the first glance. However, the feeling of obscurity will vanish if the reader invokes the concepts of ‘variable’ and ‘time’ which he or she is grown accustomed to.

Indeed, what does this definition intend to convey if translated into the language of ‘variable’ and ‘time’? To grasp this, let us consider a variable x that ranges over a given numerical sequence M , shifting from the preceding indices to the succeeding ones ... in the language of a variable and time this definition means that a ([partial]) *limit* of a numerical sequence M is such a real number a that the variable x cannot leave for ever, since the values of x become however ‘close to a ’ from time to time.”

Using the same language in infinitesimal analysis, we may express this definition in the most lucid and comprehensible manner: “If the variable x is infinitely close to a at some remote moment of time then a is a [partial] limit of M .”

Continuing the discussion of the tests of Theorem 2.3.1, we recall the following directions by Courant:

“Motivation of the rigorous definition of limit. It is no wonder that anyone who first hears the abstract definition of the limit of a sequence will fail to

comprehend it completely from the very beginning. This definition of limit arranges a game between two persons, A and B : A demands that a constant quantity a be approximately presented by some entry a_n so that the deviation be less than an arbitrary bound $\varepsilon = \varepsilon_1$ imposed exclusively by A . B meets this demand by proving the existence of such an integer $N = N_1$ that all a_n satisfy the demand of ε_1 from some a_N on.

Then A wants to demand a new smaller bound, $\varepsilon = \varepsilon_2$, while B , in turn, meets this demand by finding a new integer $N = N_2$ (possibly, much greater), and so on. If B manages always to meet the demand by A of a however small bound then we have the situation that is symbolically expressed as $a_n \rightarrow a$.

Undoubtedly, there is a psychological difficulty in mastering this rigorous definition of passage to the limit. Our mental apprehension imposes upon us the ‘dynamical’ idea of passage to the limit as a result of motion: we ‘run’ through the sequence of numbers $1, 2, 3, \dots, n, \dots$ and observe the behavior of the sequence a_n . We harbor the feeling that approximation must be visible during our ‘run.’ This ‘natural’ formulation, however, does not allow a rigorous mathematical rephrasal.

To arrive at the rigorous definition, the order of consideration should be reversed: Instead of starting with the argument n and proceeding further with the dependent variable a_n corresponding to n , we base our definition on the steps that enable us to validate the statement $a_n \rightarrow a$ successively.

Pursuing this inspection, we must start with choosing arbitrarily small interval about a , and then to check whether the condition of hitting the interval is fulfilled by assuming the variable n sufficiently large. It is in this way that we come to the rigorous definition of limit, assigning to the expressions ‘arbitrarily small bound’ and ‘sufficiently large’ n the symbolic denotations ε and N ” [64, pp. 66–67].

Compare this with the infinitesimal limit test 2.3.1(2) which reads: “If the general entry a_N is indistinguishable from a standard number a for all infinitely large N then a is declared (and, in fact, is) the limit of (a_n) .” It is beyond a doubt that the above chant successfully expresses the dynamical idea of passage to the limit.

Using the tests of Theorem 2.3.1 in the standard environment, we should always bear in mind that they apply only to standard sequences, failing in general for nonstandard “inassignable” sequences. For instance if we define the general entry as $a_n := N/n$, with $N \approx +\infty$, then $a_n \rightarrow 0$ whereas $a_N = 1$.

In other words, the tests of Theorem 2.3.1 supplement the contemporary views of limits rather than refuting or neglecting them. We may elaborate this claim as follows: Defining the concept of a convergent standard sequence, by standardization we give birth to the standard set of all convergent sequences. All in all, the conventional ε - N -definition and the nonstandard definition with actual infinites and infinitesimals coexist in a rock-solid unity.

It is also worth emphasizing that the particular applications in practice (for instance in physics) often supply us with some “actual,” “assignable”—*standard*—sequences. Moreover, these particular applications provide a definite “rigorous” meaning for an infinitely large quantity by overtly invoking or even specifying some *horizon*, i.e., an upper bound beyond which the numbers are proclaimed indiscernible. The problems of existence are also solved in practice on some semantic basis: were a physically meaningful speed absent, we would have no reasons to seek for it. Therefore, we encounter the problem of recognizing the limit of some “assignable”—*standard*—sequence.

Infinitesimal analysis gives an easy receipt: “Take a general entry of your sequence with whatever (no matter which) infinitely large index; recognize the value of this entry up to infinitesimal and, surprise, here goes the limit.” This let us comprehend better the reasons behind the infinitesimal methods of the founders of differential and integral calculus who were seeking the answers to the problems about the exact values of particular “standard” objects: areas of figures, equations of tangents to “named” curves, integrals of explicitly written analytical expressions, etc.

2.3.3. Among the most important new contribution of infinitesimal analysis we should indicate the definition of limit for a *finite sequence* $a[N] := (a_1, \dots, a_N)$, with N an infinitely large natural number. The intuitive idea behind the definition to follow happily reflects the practical tools for finding various numerical characteristics of the huge totalities of distinct entities such as thermodynamic parameters of fluid media, estimates for societal demand, etc.

2.3.4. The real a is a *microlimit* or *nearlimit value* of a finite sequence $a[N]$ provided that $a_M \approx a$ for all infinitely large M less than N . In this event we also say that $a[N]$ is *nearly convergent* to a .

If a is a limited real then the standard part ${}^\circ a$ of a is the *limit* or *S-limit* of $a[N]$. In this event we write either ${}^\circ a = \approx \lim a[N]$ or ${}^\circ a = S\text{-}\lim_{n \leq N} a_n$. So,

$${}^\circ a = \approx \lim a[N] \leftrightarrow a \in \approx \mathbb{R} \wedge (\forall M \approx +\infty, M \leq N)(a_M \approx a).$$

2.3.5. Assume that (a_n) is a standard sequence, $N \approx +\infty$, and $a \in \approx \mathbb{R}$. Then the following are equivalent:

- (1) a is a microlimit of $a[N]$;
- (2) (a_n) converges to ${}^\circ a$.

◁ The implication (2) \rightarrow (1) follows from 2.3.1(2). To prove (1) \rightarrow (2), take an arbitrary standard $\varepsilon > 0$ and put

$$A := \{m \in \mathbb{N} : (\forall n)((m \leq n \leq N) \rightarrow |a_n - {}^\circ a| \leq \varepsilon)\}.$$

The set A is nonempty since $N \in A$. Hence, A contains the least element m . If $m \approx +\infty$ then $m - 1 \approx +\infty$ and $m - 1 \in A$ by hypothesis. It follows that m is standard. Moreover, if $n \geq m$ and n is standard, then $n \leq N$ and $|a_n - {}^\circ a| \leq \varepsilon$. Hence, $(\forall \varepsilon \in {}^\circ\mathbb{R}, \varepsilon > 0)(\exists m \in {}^\circ\mathbb{N})(\forall n \in {}^\circ\mathbb{N})(n \geq m \rightarrow |a_n - {}^\circ a| \leq \varepsilon)$. By transfer, we conclude that (a_n) converges to ${}^\circ a$. \triangleright

2.3.6. The above proposition rigorously corroborates the heuristic “granted horizon principle” which prompts us to select some “physically” or “economically” actual infinite as a measure of presentability as well as a natural bound for the samples under study.

2.3.7. EXAMPLES.

$$(1) \lim_{n \rightarrow \infty} \frac{n-1}{n} = 1.$$

\triangleleft Take an infinitely large ι and note that ${}^\circ\left(\frac{\iota-1}{\iota}\right) = {}^\circ\left(1 - \frac{1}{\iota}\right) = 1$. In more detail, Euler wrote: “Since ι is an infinitely large number; therefore, $\frac{\iota-1}{\iota} = 1$. Indeed, it is clear that the greater number substitute for ι , the closer the magnitude of $\frac{\iota-1}{\iota}$ will become to one; if ι is greater than any assignable number then the fraction $\frac{\iota-1}{\iota}$ will become equal to one” [109, p. 116]. \triangleright

$$(2) \lim_{n \rightarrow \infty} \frac{n}{2^n} = 0.$$

\triangleleft If N is infinitely large then $2^N = (1+1)^N \geq N(N-1)/2$, i.e., $0 \leq N/2^N \leq 2/(N-1) \approx 0$. Hence, $N/2^N \approx 0$. \triangleright

$$(3) \lim_{n \rightarrow \infty} \sin(2\pi n!e) = 0.$$

\triangleleft Given a natural n , note that

$$0 < e - \sum_{k=1}^n \frac{1}{k!} < \frac{3}{(k+1)!}.$$

Hence,

$$0 \leq N! \left(e - \sum_{k=1}^N \frac{1}{k!} \right) \leq \frac{3N!}{(N+1)!} = \frac{3}{N+1} \approx 0$$

for all unlimited N . Put $x := 2\pi N!e$ and $y := 2\pi N! \sum_{k=1}^N 1/k!$. It follows that $x \approx y$ and so $\sin y = 0$. Obviously,

$$|\sin x - \sin y| = 2 \left| \cos \frac{x+y}{2} \sin \frac{x-y}{2} \right| \leq |x-y|,$$

which implies that $\sin x \approx y$. \triangleright

(4) Let (a_n) be a sequence such that the sequences (a_{2n}) , (a_{2n+1}) , and (a_{3n}) are convergent. Then (a_n) is convergent too.

◁ By transfer, we may assume (a_n) standard. Given an infinitely large N , observe that $2N \approx +\infty$, $2N + 1 \approx +\infty$, and $3N \approx +\infty$. Consequently, $a_{2N} \approx a$, $a_{2N+1} \approx b$, and $a_{3N} \approx c$ for some standard numbers a , b , and c . In particular, $a_{6N} \approx a \approx c$ and $a_{6N+1} \approx b \approx c$. Hence, $a = b = c$, which completes the proof. ▷

(5) If (a_n) is a vanishing sequence then

$$\lim_{n \rightarrow \infty} \frac{a_1 + \cdots + a_n}{n} = 0.$$

◁ By transfer, we may assume (a_n) standard. Take $N \approx +\infty$ and let M stand for the integral part of \sqrt{N} . Clearly, M is infinitely large. By hypothesis, $|a_N| \leq n^{-1}$ for all $n \in {}^\circ\mathbb{N}$ and so

$$\begin{aligned} s_N &:= \left| \frac{a_1 + \cdots + a_N}{N} \right| \leq \left| \frac{a_1 + \cdots + a_M}{N} \right| + \left| \frac{a_{M+1} + \cdots + a_N}{N} \right| \\ &\leq \frac{M}{N} \sup_{n \in \mathbb{N}} |a_n| + \frac{1}{n} \frac{N - M - 1}{N}. \end{aligned}$$

Since $1/N \approx 0$ and $\sup_{n \in \mathbb{N}} |a_n| \in {}^\circ\mathbb{R}$; therefore, s_N is infinitely small. ▷

(6) The space $l_\infty := l_\infty(\mathbb{N}, \mathbb{R})$ of bounded sequences admits a Banach limit; i.e., there is a continuous linear functional l on l_∞ such that

$$\begin{aligned} (\exists \lim a_n) &\rightarrow l(a) = \lim a_n; \\ \liminf a_n &\leq l(a) \leq \limsup a_n; \\ (a)(n) &:= a_{n+1} \rightarrow l(a) = l(a) \end{aligned}$$

for all members $a := (a_n)$ of l_∞ .

◁ To prove, take some infinitely large natural number N . Given a standard member $a = (a_k)$ of l_∞ , observe that

$$f(a) := \frac{1}{N} \sum_{k=N}^{2N-1} a_k$$

is limited. Indeed, since a is standard; therefore, so is the real $\|a\|_\infty := \sup_{n \in \mathbb{N}} |a_n|$. Moreover,

$$|f(a)| \leq \frac{1}{N} \sum_{k=N}^{2N-1} |a_k| \leq \frac{1}{N} \sum_{k=N}^{2N-1} \|a\|_\infty \leq \|a\|_\infty.$$

Considering that $l_\infty \times \mathbb{R}$ is a standard set, apply the standardization principle to obtain the set

$$l := * \{(a, t) \in l_\infty \times \mathbb{R} : t = {}^\circ f(a)\}.$$

Demonstrate that l is a sought Banach limit. To this end, check first that l is a function. In particular, we have to show that

$$(\forall a \in l_\infty)(\forall t_1, t_2 \in \mathbb{R})((a, t_1) \in l \wedge (a, t_2) \in l \rightarrow t_1 = t_2).$$

By transfer, we may assume a , t_1 , and t_2 standard. In this event $t_1 = {}^\circ f(a)$ and $t_2 = {}^\circ f(a)$ by the definition of standardization. We are done on recalling that the standard part of a real is defined uniquely by 2.2.16.

The linearity of l may be proven in much the same way. It is also obvious that $a \geq 0 \rightarrow l(a) \geq 0$; in other words, l is a positive linear functional.

Let a be a standard sequence converging to \bar{a} . Given a standard $\varepsilon > 0$, we then infer from 2.3.1(2) that $|a_N - \bar{a}| \leq \varepsilon, \dots, |a_{2N-1} - \bar{a}| \leq \varepsilon$ since a_M is infinitely close to \bar{a} for all $M \geq N$. Hence,

$$|f(a) - \bar{a}| = \left| \frac{1}{N} \sum_{k=N}^{2N-1} (a_k - \bar{a}) \right| \leq \varepsilon,$$

i.e., $\bar{a} = {}^\circ f(a)$. This, together with the positivity of l , provides the sought estimates.

We are left with establishing that l is shift-invariant, i.e., $l('a) = l(a)$ for all $a \in l_\infty$. To this end, we again assume a standard, which implies that so is $'a$ and, consequently,

$$\begin{aligned} l('a) &= {}^\circ \left(\frac{1}{N} \sum_{k=N}^{2N-1} a_{k+1} \right) \text{st} (N^{-1}(a_{N+1} + a_{N+2} + \dots + a_{2N})) \\ &= \text{st} \left(\frac{1}{N} \sum_{k=N}^{2N-1} a_k + \frac{1}{N} a_{2N} - \frac{1}{N} a_N \right) = {}^\circ (f(a) + N^{-1} a_{2N} - N^{-1} a_N) \\ &= {}^\circ f(a) + (N^{-1} a_{2N}) - {}^\circ (N^{-1} a_N) = {}^\circ f(a) = l(a). \end{aligned}$$

The above derivation uses the fact that a_{2N}/N and a_N/N are limited as well as the properties of the standard part operation (cf. 2.2.18). \triangleright

2.3.8. Theorem. *Let f be a standard real function and let x be a standard point of $\text{dom}(f)$, the domain of f . Then the following are equivalent:*

- (1) f is continuous at x ;
- (2) f sends each point infinitely close to x to some point infinitely close to $f(x)$; in symbols,

$$x' \approx x, x' \in \text{dom}(f) \rightarrow f(x') \approx f(x).$$

◁ (1) → (2): Let $\varepsilon > 0$ be a standard real. There is some $\delta > 0$ such that $|f(x') - f(x)| \leq \varepsilon$ whenever $|x' - x| \leq \delta$ and $x' \in \text{dom}(f)$. By transfer, there is a standard δ enjoying the same property. If $x' \approx x$ and $x' \in \text{dom}(f)$ then, obviously, $|x' - x| \leq \delta$ (since $\delta \in {}^\circ\mathbb{R}$). Hence, $|f(x) - f(x')| \leq \varepsilon$. Since $\varepsilon \in {}^\circ\mathbb{R}$ is arbitrary; therefore, $f(x') \approx f(x)$.

(2) → (1): Take an arbitrary $\varepsilon > 0$. We have to find some δ fitting the “ ε - δ -definition.” By transfer, it suffices to do this on assuming ε standard. However, in case ε is standard we may take as δ whatever strictly positive infinitesimal. ▷

2.3.9. In view of 2.3.8(2), we call a function $f : \text{dom}(f) \rightarrow \mathbb{R}$ *microcontinuous* at x provided that $x \in \text{dom}(f)$ and $f(x') \approx f(x)$ whenever $x' \in \text{dom}(f)$ and $x' \approx x$.

2.3.10. Commenting on Theorem 2.3.8, we note that microcontinuity and ε - δ -continuity at a point mean the same for a standard function at a standard point. We may also repeat the discussion of 2.3.2 and emphasize, following Courant, that “as in the case of the limit of a sequence, the Cauchy definition that rests, so to say, on the reversal of the intuitively acceptable order in which we would prefer to consider variables. Instead of starting from the independent variable and passing then to the dependent variable, we first direct our attention to the ‘accuracy estimate’ ε , and then we try to restrict the corresponding ‘arena’ of δ for the independent variable” [64, p. 73].

Theorem 2.3.8 eliminates the unpleasant reversal of quantifiers for all standard functions and points. It is curious as well as luminous to observe also that Courant referred to the ε - δ -definition of continuity as the Cauchy definition whereas it is the definition of microcontinuity that mimics the words “an infinitely small increment given to the variable produces an infinitely small increment of the function itself” (cf. 1.5).

At the same time, the ε - δ -definition of continuity, applicable to all functions with no exception, is only implicitly reconstructible from microcontinuity at a point by standardization. As always, the standard and nonstandard treatments reveal their intricate but genuine unity in regard to continuity. This demonstrates that the new concept of microcontinuity of a function at a point is a valuable piece of mathematical acquisition. The following propositions will widen the scope of our understanding of microcontinuity.

2.3.11. EXAMPLES.

(1) *The function $x \mapsto x^2$ is microcontinuous at no unlimited point $t \in \mathbb{R}$.*

◁ Indeed, $t + t^{-1} \approx t$ and, at the same time, $(t + t^{-1})^2 - t^2 \approx 2$. ▷

(2) *Let δ be a strictly positive infinitesimal. Consider the function $x \mapsto \delta \sin x^{-1}$ defined as zero at the origin. This function is discontinuous at zero but microcontinuous.*

◁ It suffices to note that $\sin x \in {}^\approx\mathbb{R}$ for $x \in \mathbb{R}$ and refer to the properties of infinitesimals in 2.2.8. ▷

2.3.12. Uniform Continuity Test. *If f is a standard real function f then*

(1) *f is microcontinuous, i.e., f is microcontinuous at every point of $\text{dom}(f)$ or, in symbols,*

$$(\forall x, x' \in \text{dom}(f))(x' \approx x \rightarrow f(x') \approx f(x));$$

(2) *f is uniformly continuous.*

◁ (1) \rightarrow (2): Let $\varepsilon > 0$ be standard, and let $\delta > 0$ be infinitesimal. Obviously, $x \approx x' \rightarrow |x - x'| \leq \delta$. Therefore,

$$(\forall \varepsilon \in {}^\circ\mathbb{R}, \varepsilon > 0)(\exists \delta > 0)(\forall x, x' \in \text{dom}(f)) \\ (|x - x'| \leq \delta \rightarrow |f(x) - f(x')| \leq \varepsilon).$$

By transfer we conclude that f is uniformly continuous.

(2) \rightarrow (1): By transfer, to each standard $\varepsilon > 0$ there is some standard $\delta > 0$ satisfying $|x - x'| \leq \delta \rightarrow |f(x) - f(x')| \leq \varepsilon$ for all $x, x' \in \text{dom}(f)$. Observing that $x \approx x' \rightarrow |x - x'| \leq \delta$, we complete the proof. ▷

2.3.13. *Let f be a standard function given in a standard neighborhood of a standard point x of ${}^\circ\mathbb{R}$. The following are equivalent:*

- (1) *f is differentiable at x and $f'(x) = t$;*
 (2) *If h is a nonzero infinitesimal then*

$$t = \text{st}((f(x+h) - f(x))/h).$$

◁ This is straightforward from 2.3.8. ▷

2.3.14. *Assume that y is a standard function given in a neighborhood of a standard point x and differentiable at x . Let dx be an arbitrary nonzero infinitesimal. Denote (in the wake of Leibniz) by dy the differential of y at x calculated at dx . Then*

$$dy \approx y(x+dx) - y(x), \quad \frac{dy}{dx} = y'(x).$$

◁ By the definition of Leibniz (cf. 1.1), from 2.3.9 we infer that

$$dy = y'(x)dx, \quad y'(x) = \text{st}\left(\frac{y(x+dx) - y(x)}{dx}\right).$$

Hence,

$$dy \approx \frac{y(x+dx) - y(x)}{dx}dx = y(x+dx) - y(x),$$

which proves the first claim. The second follows from 2.3.10. ▷

2.3.15. The “nonstandard” treatments of the role of infinitesimals in the definitions of derivatives, differentials, and increments as given in 2.3.13 and 2.3.14, supplement the following directions by Euler:

“I have already remarked that in differential calculus the problem of finding differentials should be understood in a relative rather than absolute sense. This means that if y is a function of x then it is not its differential but rather the ratio of this differential to the differential dx that should be determined. Indeed, since all differentials are exactly equal to zero; therefore, whatever the function y of the quantity x might be, dy is always equal to zero; therefore, in the absolute sense there is nothing more to be sought for.

The sound statement of the problem is as follows: x takes an infinitely small, i.e., vanishing increment dx [= *evanescens*, an actual number that ‘is exactly equal to zero’]; the problem is to determine the ratio of the resulting increment of the function y to dx s. In fact, both increments = 0; however, there is some finite ratio between them that is perfectly revealed in differential calculus.

For instance, if $y = x^2$ then, as demonstrated in differential calculus, $\frac{dy}{dx} = 2x$ and this ratio of the increments is valid only if the increment dx which generates dy is considered to be exactly equal to zero. Nevertheless, after this warning on the true notion of differential has been made, it is allowed to use the conventional expressions that treat the differentials as in the absolute sense provided though that the truth be constantly borne in mind. For instance, we have the right to say: if $y = x^2$ then $dy = 2xdx$. In actuality, if somebody said that $dy = 3xdx$ or that $dy = 4xdx$ then this would be not false either, since even these equalities are valid as $dx = 0$ and $dy = 0$. Only the first equality, however, agrees with the true relation $\frac{dy}{dx} = 2x$ [111, p. 9].

The reader will note that Euler used the sign “=” at the places where we write “ \approx ” (see 2.2.10). Moreover, we may emphasize that he was seeking for the differentials that he presumed existent, while working with particular (differentiable) functions. In these circumstances, it would be quite legitimate to use for finding the differential any infinitely small dx chosen in any way.

Therefore, Euler had full grounds to say that the differential dy (calculated at an infinitely small dx) “is exactly equal to zero”; the differential dy is exactly the increment, i.e., the “absolute differential”; and at the same time the differential dy is the “fourth proportional” for infinitely small increments, i.e., in our notation:

$$\begin{aligned} \circ dy &= 0, & \circ(dy - (y(x + dx) - y(x))) &= 0; \\ \circ\left(\frac{dy}{dx} - \frac{y(x + dx) - y(x)}{dx}\right) &= 0. \end{aligned}$$

The above analysis demonstrates the soundness of the ideas and views of Euler

in handling “assignable”—standard—objects, the function y and the point x , on assuming the crucial hypothesis that dx is infinitely small.

In the limelight of the above arguments we must exercise due criticism at pondering over the following words by Courant:

“If we want to understand the essence of differential calculus then we should beware of viewing a derivative as a ratio of two actually existing ‘infinitely small quantities.’ The point is that we must always start with arranging the ratio of the increments $\Delta y/\Delta x$, where the difference Δx is not equal to zero. Then we should imagine that, either by way of transforming this ratio or by some other way, passage to the limit has been accomplished. But in no case you can imagine that at *first* there is some transition from Δx to an infinitely small quantity dx which is still other than zero and from Δy to dy after which we divide these ‘infinitely small quantities’ one by the other. Such a view of the derivative is absolutely incompatible with the requirement of the mathematical clarity of notions, and it makes hardly any sense at all” [64, pp. 126–128].

The excessive stiffness of the last phrase is but partially smoothed down by the further clarification:

“A physicist, a biologist, a technician, or any other specialist who has to deal with these notations in practice has therefore the right to identify, within the accuracy required, the derivative with the ratio of the increments... . ‘Physically infinitesimal’ quantities have an exact sense. They are undoubtedly finite nonzero quantities, chosen though to be sufficiently small in the problem under discussion, less for instance than some fraction of a wavelength or smaller than the distance between two electrons in an atom; etc., generally speaking, less than some desirable degree of accuracy” [64, p. 135].

2.3.16. We proceed with discussion of the basics of integral calculus and start with the “infinitesimal” definition of the Riemann integral.

2.3.17. Let $f : [a, b] \rightarrow \mathbb{R}$ be a standard continuous function, and let $a = x_1 < x_2 < \dots < x_N < x_{N+1} = b$ be a partition of $[a, b]$ satisfying $\xi_k \in [x_k, x_{k+1}]$ and $x_k \approx x_{k+1}$ for $k := 1, \dots, N$. Then

$$\int_a^b f(x)dx = \circ \left(\sum_{k=1}^N f(\xi_k)(x_{k+1} - x_k) \right).$$

◁ Observe first that N is infinitely large, and appeal to the conventional definition of the integral, as well as Theorems 2.3.1 and 2.3.12 for f . ▷

2.3.18. Theorem (the basic principle of integral calculus). “In calculation of the sum of an infinitely large number of infinitely small addends (of the same sign) it is possible to subtract from each addend a higher-order infinitesimal.”

◁ Consider the sum $\circ \sum_{k=1}^N \alpha_k = t$, with $\alpha_k \approx 0$. By hypothesis, $\beta_k := \alpha_k - o(\alpha_k)$. By 2.2.13 (2) $\beta_k \sim \alpha_k$, and so

$$t = \circ \left(\sum_{k=1}^N \alpha_k \right) = \circ \left(\sum_{k=1}^N \beta_k \right),$$

which completes the proof. ▷

2.3.19. The above propositions give formal grounds to consider the integral as a finite sum of infinitely small terms, i.e., they justify the antique view of integration as a particular instance of summation.

In this respect we find it illuminating to quote the following definition of integral (“with a variable upper limit”) by Euler:

“Integration is usually defined as follows: It is said to be summation of all values of the differential expression $X dx$ which the variable x assumes consecutively at all values that differ by the difference dx starting from some given value up to x , with this difference taken to be infinitely small. ... From the method presented it is in any case clear that integration results from summation to within any accuracy, whereas integration cannot be accomplished exactly in any other way than by assuming these differences to be infinitesimals, i.e., zeros” [111, p. 163].

We are happy to note once again that in order to find the integral of a standard continuous function it suffices, as follows from the facts presented above, to calculate the “exact value” or standard part of only one finite sum of an infinitely large number of infinitely small addends in which we may neglect all higher-order infinitesimals. This technique does not work in general for arbitrary nonstandard functions. In other words, we again discover, as has been repeatedly the case before, that the nonstandard views of the objects of mathematical analysis supplement, refine, and elaborate but never refute their conventional counterparts.

2.3.20. All these facts manifest that the nonstandard methods of today are direct descendants of the calculus of infinitesimals. That is why the term “infinitesimal analysis” regains popularity as better reflecting the essence of the matter than a somewhat extravagant term “nonstandard analysis” often causing irritation, which is quite understandable in the long run.

It deserves special emphasis that the idea of actual infinites and infinitesimals has never abandoned the toolkit of natural sciences but rather took a short leave from mathematics for about thirty years. This saves us from expatiating eloquently on the versatility and importance of infinitesimal analysis.

Chapter 3

Set-Theoretic Formalisms of Infinitesimal Analysis

The naive discussion of Chapter 2 of the qualitative differences between the standard (explicit) and nonstandard (implicit) methods of introducing mathematical objects has enriched our views of the infinitely large and infinitely small, making them more rigorous but still agreeable with intuition. Our most valuable acquisition was a better understanding of the methods of reasoning which are deeply rooted in the bygones of differential and integral calculus.

At the same time, we face serious complications even in simple examples. First of all, the criteria remain obscure for distinguishing between standard and nonstandard entities, which stirs up a premonition that the principles of infinitesimal analysis might be mistreated. Discontent is aggravated with appearance of the objects that, on the one hand, seem quite legitimate as constructed with the routine mathematical tools but, on the other hand, cannot achieve the status of conventional sets without contradiction.

These are galore: various monads, proximities, the collections of standard parts of infinite subsets of the reals, O 's and o 's, etc. Still more annoying is the fact that the “mathematical law” $x \mapsto \circ x$, the standard part operation acting from $\overline{\mathbb{R}}$ to $\overline{\mathbb{R}}$, is not a function. The point is that the concept of function had been established in mathematics long before the formulation of the set-theoretic stance.

It was as far back as in 1755 that Euler wrote: “If some quantities depend on the others in such a way that when the latter are varied the former are also subject to varying, then the former are called the functions of the latter. This denomination has an extremely wide range of application; it embraces all manners in which one quantity may be determined from the others. Thus, ... all quantities that depend on x in whatever manner, i.e., that are determined by x , are called the functions of x ” [110, p. 38].

This dynamical idea of how one object may depend on the other is not conveyed

completely by the “stationary” set-theoretic view of a function as a set which prevails nowadays. “It is a formal set-theoretic model of the intuitive idea of a function, a *model* that captures an aspect of the idea, but not its full significance” [133, p. 20].

We recall in this regard that if $s, t \in [0, 1]$ then

$${}^\circ(s + t) = {}^\circ s + {}^\circ t, \quad {}^\circ 0 = 0, \quad {}^\circ 1 = 1,$$

and, moreover, ${}^\circ t = 0$ for all t in some interval $[0, h]$, where h is a strictly positive real (every nonzero positive infinitesimal will do). The presence of such a “numerical” function is an outright contradiction or, to put it mildly, a harbinger of antinomy.

These circumstances call for clarifying, immediately and explicitly, the concepts and means we use as well as specifying the foundations we rest them on.

As we have already mentioned, infinitesimal analysis acquires justification within the set-theoretic stance. More exactly, it appears that the ideas of the naive nonstandard set theory we have presented above can be placed on the same (and so, equally solid) foundations as cantorion set theory or, strictly speaking, the axiomatic set theories “approximating the latter from below.”

In order to bring into focus the relations between mathematical analysis and set theory, the following statements are worth comparing:

Analysis ... is the science of the infinite itself.

LEIBNIZ

Mathematical analysis is just the science of the infinite.

This old definition lives through ages.

LUZIN

SET THEORY, an area of mathematics which studies the general properties of sets, primarily, of infinite sets.

THE GREAT ENCYCLOPEDIA DICTIONARY

Consequently, the very notion of the infinite intertwines analysis and set theory quite tightly. At the same time we should never forget that the classical articles by Cantor appeared two centuries after the invention of calculus.

The attempt at grounding mathematics on set theory could be compared with a modern method of building erection, rack mounting, when a house is assembled starting with upper stores, “from attic to cellar.” By the way, this technology requires that the footing of the building to be erected has been laid before the rack mounting begins. Likewise, the initial footing of mathematical analysis is a product of the material and mental activities of mankind.

The present-day mathematics leans its basic parts on set theory. In other words, the set-theoretic foundation has been floated under the “living quarters” of mathematics. Only the future will reveal what is going to happen next. By

now we may just state that the process continues of erecting the edifice of future mathematics and that this process is fraught with drastic changes.

Aggravation of the state of the art, collision of opinions, and a fierce struggle of ideas are faithful witnesses of rapid development.

A collection of quotations to follow (far from claiming for completeness) will illustrate the process of polarization of views now in progress.

PRO

After an initial period of distrust the newly created set theory made a triumphal inroad in all fields of mathematics. Its influence on mathematics of the present century is clearly visible in the choice of modern problems and in the way these problems are solved. Applications of set theory are thus immense.

KURATOWSKI AND
MOSTOWSKI [254, p. v]

Part of the creation of Georg Cantor is, of course, set theory, and some of this is now taught in high school and earlier. This is another of the domains of mathematics that many persons thought could never be of the remotest practical use, and how wrong they were. Elementary sets even find their application in little collections of murder mysteries. Set theory has well-known connections with computer programs and these affect an untold number of practical projects.

YOUNG [533, p. 102]

Concluding the preliminary discussion we emphasize that only now, after dispelling the illusion that it is possible to provide some final “absolute” foundation for infinitesimal analysis (as well as for the whole of mathematics) by the set-theoretic or whatever stance, we may proceed with exposing some available implementations of this project.

3.1. The Language of Set Theory

Axiomatic set theories are bylaws of sound set formation. In evocative words, every axiomatics of set theory describes a world or universe that consists of all sets we need for adequate expression of our intuitive conception of the treasure-trove of the “cantorian paradise,” the all-embracing universe of the naive set theory.

CONTRA

It is claimed that the theory of sets is important for the progress of science and technology, while presenting one of the most recent achievements in mathematics. In actuality, the theory of sets has nothing to do with the progress of science and technology nor it is one of the most recent achievements of mathematics.

PONTRYAGIN [400, p. 6]

Mathematics, based on Cantor set theory, changed to mathematics of Cantor set theory. ... Contemporary mathematics thus studies a construction whose relation to the real world is at least problematic. ... This makes the role of mathematics as a scientific and useful method rather questionable. Mathematics can be degraded to a mere game played in some specific artificial world. This is not a danger for mathematics in the future but an immediate crisis of contemporary mathematics.

VOPĚNKA [513]

The present-day mathematics customarily expounds and studies any attractive axiomatics as a formal theory. We acknowledge readily that a formal approach has proven itself to be exceptionally productive and successful in spite of its obvious limitations stemming from the fact that mathematics reduces only in part to the syntax of mathematical texts. This success is in many respects due to the paucity of formal means since the semiotic aspects, if properly distinguished, invoke the insurmountable problem of meaning. The list of achievements of the formal approach contains the celebrated Gödel completeness and incompleteness theorems, independence of the continuum hypothesis and of the axiom of choice, Boolean valued analysis, etc.

The cornerstone of a formal theory is its language. Intending to give the latter some exact description and to study the properties of the theory, we are impelled to use another language that differs in general from the original language. It is in common parlance to call this extra language the *metalanguage* of the theory in question. The metalanguage mostly presents a collection of fragments of natural languages trimmed and formalized slightly but enriched lavishly with numerous technical terms. The tools of the metalanguage of a theory are of utmost importance for metamathematics. Since we are interested in applicable rather than metamathematical aspects of an axiomatic set theory, we never impose extremely stringent constraints on the metalanguage of the theory. In particular, we use the expressive means and level of rigor that are common to the mathematical routine.

3.1.1. Each axiomatic set theory is a *formal system*. The ingredients of such a theory are its alphabet, formulas, axioms, and rules of inference. The *alphabet* of a formal theory is a distinguished set A of symbols of an arbitrary nature, i.e., a cantorion set of letters. Finite sequences of letters of A , possibly with blanks, are called *expressions*, or *records*, or *texts*. If we distinguish some set $\Phi(A)$ of expressions by giving detailed prescriptions, algorithms, etc.; then we declare given a language with alphabet A and call the chosen expressions *well-formed formulas*. The next step consists in selecting some finite (or infinite) families of formulas called *axioms* in company with explicit description of the admissible *rules of inference* which might be viewed as abstract relations on $\Phi(A)$. A *theorem* is a formula resulting from the axioms by successive application of finitely many rules of inference. Using common parlance, we express this in a freer and cozier fashion as follows: The theorems of a formal theory comprise the least set of formulas which contains all axioms and is closed under the rules of inference of the theory.

3.1.2. Of primary interest for us is some special formal language, the so-called *first-order language* of predicate calculus.

Recall that the *signature* σ of a language is a triple (F, P, a) where F and P are some sets called the set of *function* or *operation symbols* and the set of *predicate symbols*, respectively, while a is a mapping of $F \cup P$ into the set of natural numbers.

Say that $u \in F \cup P$ is an n -ary symbol or n -place symbol whenever $a(u) = n$. Regarding the *alphabet* of a first-order language of signature σ , we usually distinguish:

- (1) the set of *symbols* of signature σ , i.e., the set $F \cup P$;
- (2) the set of *variables* composed of lower case or upper case Latin letters possibly with indices;
- (3) the set of *propositional connectives*: \wedge , conjunction; \vee , disjunction; \rightarrow , implication; and \neg , negation;
- (4) the set of the symbols of *quantifiers*: \forall , the symbol of a universal quantifier, and \exists , the symbol of an existential quantifier;
- (5) the sign of *equality* $=$;
- (6) the set of *auxiliary symbols*: the opening parenthesis $($; the closing parenthesis $)$; and the comma $,$.

3.1.3. In the language of set theory we distinguish terms and formulas.

(1) A *term* of signature σ is an element of the least set of expressions of the language (of the same signature σ) obeying the following conditions:

- (a) Each variable is a term;
- (b) Each nullary function symbol is a term;
- (c) If $f \in F$, $a(f) = n$, and t_1, \dots, t_n are terms then $f(t_1, \dots, t_n)$ is a term.

(2) An *atomic formula* of signature σ is an expression of the kind

$$t_1 = t_2, \quad p(y_1, \dots, y_n), \quad q,$$

where $t_1, t_2, y_1, \dots, y_n$ are terms of signature σ , the letter p stands for some n -ary predicate symbol, and q is a nullary predicate symbol.

(3) The *formulas of signature* σ constitute the least set of records obeying the following conditions:

- (a) Each atomic formula of signature σ is a formula of signature σ ;
- (b) If φ and ψ are formulas of signature σ then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $\neg\varphi$ are formulas of signature σ , too;
- (c) If φ is a formula of signature σ and x is a variable then $(\forall x)\varphi$ and $(\exists x)\varphi$ are formulas of signature σ too.

A variable x is *bound* in some formula φ or belongs to the domain of a quantifier provided that x appears in a subformula of φ of the kind $(\forall x)\psi$ or $(\exists x)\psi$. In the opposite case, x is *unbound* or *free* in φ . We also speak about *free* or *bound occurrence* of a variable in a formula. Intending to stress that only the variables x_1, \dots, x_n are unbound in the formula φ , we write $\varphi = \varphi(x_1, \dots, x_n)$, or simply $\varphi(x_1, \dots, x_n)$. The words “proposition” and “statement” are informally treated

as synonyms of “formula.” A formula with no unbound variables is a *sentence*. Speaking about *verity* or *falsity* of φ , we imply the *universal closure* of φ which results from generalization of φ by every bound variable of φ . It is also worth observing that quantification is admissible only by variables. In fact, the words “first-order” distinguish this syntactic feature of the formal languages we discuss.

3.1.4. The *language of set theory* is a first-order language whose signature contains only one binary predicate symbol \in and so it has no predicates but \in nor any function symbols. So, set theory is a simple instance of the abstract *first-order theories*. We agree to write $x \in y$ instead of $\in(x, y)$ and say that x is an *element* or a *member* of y . It is also in common parlance to speak of *membership* or *containment*. As usual, a formula of set theory is a formal text resulting from the atomic formulas like $x \in y$ and $x = y$ by appropriate usage of propositional connectives and quantifiers.

Set theory (or strictly speaking, *the* set theory we profess in this book) bases upon the laws of classical logic. In other words, set theory uses the common logical axioms and rules of inference of predicate calculus which are listed in nearly every manual on mathematical logic (see, for instance, [105, 238, 444]). Note also that the instance of predicate calculus we use in this book appears often with some of the epithets *classical*, or *lower*, or *narrow*, or *first-order* and is formally addressed as the *first-order classical predicate calculus with equality*. In addition, a particular set theory contains some special nonlogical axioms that legitimize the conceptions of sets and classes we want to explicate. By reasonably varying the special axioms, we may come to axiomatic set theories that differ in the power of expression. This section describes one of the most popular axiomatic set theories, *Zermelo–Fraenkel set theory* symbolized as ZF or ZFC if the axiom of choice is available or stressed.

3.1.5. Among the best conveniences of any metalanguage we must mention abbreviations. The point is that formalization of the simplest fragments of workable mathematics leads to bulky texts whose recording and playing back is problematic for both physical and psychological reasons. For that reason we must introduce many abbreviations, building a more convenient abridged dialect of the initial symbolic language. Naturally, this is reasonable only if we ensure a principal possibility of unambiguous translation from the dialect to the original and vice versa. In accord with our intentions, we will not expatiate on the exact technique of abbreviation and translation and adhere to the every-day practice of doing Math. For instance, we use the *assignment operator* or *definor* $:=$ throughout the book, with no fuss about accompanying formal subtleties.

3.1.6. We now give some examples of abbreviated texts in the language of set theory. These examples rely on the intuition of the naive set theory. We start with

the most customary instances. Here they are:

$$\begin{aligned} (\exists! x) \varphi(x) &:= (\exists x) \varphi(x) \wedge (\forall x)(\forall y)(\varphi(x) \wedge \varphi(y) \rightarrow x = y); \\ (\exists x \in y) \varphi &:= (\exists x)(x \in y \wedge \varphi); \\ (\forall x \in y) \varphi &:= (\forall x)(x \in y \rightarrow \varphi), \end{aligned}$$

with φ a formula. As usual, we put $x \neq y := \neg(x = y)$ and $x \notin y := \neg(x \in y)$. Also, we use the routine conventions about the traditional operations on sets:

$$\begin{aligned} x \subset y &:= (\forall z)(z \in x \rightarrow z \in y); \\ u = \bigcup x &:= (\forall z)(z \in u \leftrightarrow (\exists y \in x)z \in y); \\ u = \bigcap x &:= (\forall z)(z \in u \leftrightarrow (\forall y \in x)z \in y); \\ u = y - x := y \setminus x &:= (\forall z)(z \in u \leftrightarrow (z \in y \wedge z \notin x)). \end{aligned}$$

Given a formula φ , we introduce the collection $\mathcal{P}_\varphi(x)$ of all subsets of x which satisfies φ as follows

$$u = \mathcal{P}_\varphi(x) := (\forall z)(z \in u \leftrightarrow (z \subset x) \wedge \varphi(z)).$$

In particular, if $\text{fin}(y)$ means that y is a finite set then $\mathcal{P}_{\text{fin}}(x)$ comprises all finite subsets of x . We call a set u *empty* and denote it by \emptyset if u contains no elements. In other words,

$$u = \emptyset := (\forall x)(x \in u \leftrightarrow x \neq x).$$

An empty set is unique practically in every set theory and so we refer to \emptyset as *the* empty set.

These examples use one of the commonest methods of abbreviation, namely, omission of some parentheses.

3.1.7. The statement that x is the *unordered pair* of elements y and z is formalized as follows:

$$(\forall u)(u \in x \leftrightarrow u = y \vee u = z).$$

In this event we put $\{y, z\} := x$ and speak about the *unordered pair* $\{y, z\}$. Note that braces do not belong to the original alphabet and so they are *metasymbols*, i.e. symbols of the metalanguage.

An *ordered pair* or *ordered couple* and an *ordered n -tuple* result from the Kuratowski trick:

$$\begin{aligned} (x, y) &:= \langle x, y \rangle := \{\{x\}, \{x, y\}\}; \\ (x_1, \dots, x_n) &:= \langle x_1, \dots, x_n \rangle := \langle \langle x_1, \dots, x_{n-1} \rangle, x_n \rangle, \end{aligned}$$

where $\{x\} := \{x, x\}$. Observe the overuse of parentheses. This is inevitable and must never be regarded as pretext for introducing new symbols. Also note that it is in common parlance to omit the attribute “ordered” while speaking about pairs, couples, and tuples.

The agreements we made enable us to ascribe a formal meaning to the expression “ X is the *cartesian product* $Y \times Z$ of Y and Z .” Namely, we put $X := \{(y, z) : y \in Y, z \in Z\}$. Note also that the nickname “product” is in common parlance for “cartesian product.”

3.1.8. Consider the following propositions:

- (1) $\text{Rel}(X)$;
- (2) $Y = \text{dom}(X)$;
- (3) $Z = \text{im}(X)$.

Formally these read as follows:

- (1') $(\forall u)(u \in X \rightarrow (\exists v)(\exists w) u = (v, w))$;
- (2') $(\forall u)(u \in Y \leftrightarrow (\exists v)(\exists w) w = (u, v) \wedge w \in X)$;
- (3') $(\forall u)(u \in Z \leftrightarrow (\exists v)(\exists w) w = (v, u) \wedge w \in X)$.

In other words, we state in (1)–(3) that the members of X are couples, Y is the collection of the first coordinates of the members of X , and Z comprises the second coordinates of the members of X . It is in common parlance to say that Y is the *domain* of X , and Z is the *range* or *image* of X . In this event we refer to X as an *abstract relation*.

We express the fact that X is *single-valued* or, in symbols, $\text{Un}(X)$ by the formula

$$\text{Un}(X) := (\forall u)(\forall v_1)(\forall v_2)((u, v_1) \in X \wedge (u, v_2) \in X \rightarrow v_1 = v_2).$$

We put $\text{Fnc}(X) := \text{Func}(X) := \text{Un}(X) \wedge \text{Rel}(X)$. In case $\text{Fnc}(X)$ is valid, we have many obvious reasons to call X a *function* or even a *function-class*. Naturally, this implies that we will sometimes use the term “function-set” emphasizing the cases in which X is a function and set simultaneously. Paraphrasing the membership $(u, v) \in X$, we write $v = X(u)$, $X : u \mapsto v$, etc. We say that F is a *mapping* or *function* from X to Y , implying that every member of F belongs to $X \times Y$, while F is single-valued, and the domain of F coincides with X ; that is,

$$F : X \rightarrow Y := F \subset X \times Y \wedge \text{Fnc}(F) \wedge \text{dom}(F) = X.$$

The term *function-class* is also applied to F if we want to stress that F is a class.

The *restriction* of X to U is by definition $X \cap (U \times \text{im}(X))$. We denote it by $X \upharpoonright U$, or $X|U$, or $X|_U$. If there is a unique z satisfying $(y, z) \in X$ then we put

$X \ulcorner y := z$. We finally let $X \llcorner y := \text{im}(X \upharpoonright y)$. Instead of $X \ulcorner \{y\}$ we write $X(y)$ or even Xy when this does not lead to misunderstanding. It is worth emphasizing that we always exercise a liberal view on placing and removing parentheses. In other words, we insert or eliminate parentheses, influenced as a rule by what is convenient or needed for formal presentation of a record we discuss.

Abstract relations deserve special attention. Relevant details follow.

A *correspondence* Φ from X to Y is a triple $\Phi := (F, X, Y)$ where F is some subset of the product $X \times Y$. Clearly, $\text{Rel}(F)$ holds. It is in common parlance to say that F is the *graph* of Φ , in symbols, $\text{Gr}(\Phi) = F$; while X is the *domain of departure* and Y is the *domain of arrival* or *target* of Φ . Recall that a *relation* or a *binary relation* on X is a correspondence whose domain of departure and target are the same set X .

The *image* of $A \subset X$ under Φ is the projection of $(A \times Y) \cap F$ to Y . The image of A under F is denoted by $\Phi(A)$ or simply $F(A)$. Thus,

$$\Phi(A) := F(A) := \{y \in Y : (\exists x \in A)((x, y) \in F)\}.$$

To define a correspondence Φ amounts to describing the mapping

$$\tilde{\Phi} : x \mapsto \Phi(\{x\}) \in \mathcal{P}(Y) \quad (x \in X),$$

where $\mathcal{P}(Y)$ stands for the *powerset* or *boolean* of Y which is the collection of all subsets of Y . Abusing the language, we often identify the mapping $\tilde{\Phi}$, the correspondence Φ , and the graph of Φ , denoting these three objects by the same letter. We also write $\Phi(x)$ instead of $\Phi(\{x\})$.

The *domain of definition* or simply *domain* of Φ is the domain of the graph of Φ . In other words,

$$\text{dom}(\Phi) := \{x \in X : \Phi(x) \neq \emptyset\}.$$

By analogy, the image of a correspondence is the image of its graph.

3.1.9. Assume that X and Y are abstract relations; i.e., $\text{Rel}(X)$ and $\text{Rel}(Y)$. We may arrange the *composite* of X and Y , denoted by the symbol $Y \circ X$, collecting all couples (x, z) such that $(x, y) \in X$ and $(y, z) \in Y$ for some y :

$$(\forall u)(u \in Y \circ X \leftrightarrow (\exists x)(\exists y)(\exists z)(x, y) \in X \wedge (y, z) \in Y \wedge u = (x, z)).$$

The *inverse* of X , in symbols X^{-1} , is defined as

$$(\forall u)(u \in X^{-1} \leftrightarrow (\exists x)(\exists y)(x, y) \in X \wedge u = (y, x)).$$

The symbol I_X denotes the *identity relation* or *identity mapping* on X , i.e.,

$$(\forall u)(u \in I_X \leftrightarrow (\exists x)(x \in X \wedge u = (x, x))).$$

We will elaborate the above for correspondences. To this end, assume that $\Phi := (F, X, Y)$ is a correspondence from X to Y . Assign $F^{-1} := \{(y, x) \in Y \times X : (x, y) \in F\}$. The correspondence $\Phi^{-1} := (F^{-1}, Y, X)$ is the *inverse* of Φ . Consider another correspondence $\Psi := (G, Y, Z)$. Denote by H the image of $(F \times Z) \cap (X \times G)$ under the mapping $(x, y, z) \mapsto (x, z)$. Clearly,

$$H = \{(x, z) \in X \times Z : (\exists y \in Y)((x, y) \in F \wedge (y, z) \in G)\}.$$

Hence, H coincides with the composite $G \circ F$ of the graphs G and F . The correspondence $\Psi \circ \Phi := (G \circ F, X, Z)$ is the *composite*, or *composition*, or *superposition* of Φ and Ψ . We have the following obvious equalities:

$$(\Psi \circ \Phi)^{-1} = \Phi^{-1} \circ \Psi^{-1}, \quad \Theta \circ (\Psi \circ \Phi) = (\Theta \circ \Psi) \circ \Phi.$$

A few words about another abbreviation related to correspondences: Consider $\Phi := (F, X, Y)$. The *polar* $\pi_\Phi(A)$ of $A \subset X$ under Φ is the collection of all $y \in Y$ satisfying $A \times \{y\} \subset F$. In other words,

$$\pi_\Phi(A) := \pi_F(A) := \{y \in Y : (\forall x \in A)((x, y) \in F)\}.$$

If Φ is fixed then we abbreviate $\pi_\Phi(A)$ to $\pi(A)$ and $\pi_{\Phi^{-1}}(A)$ to $\pi^{-1}(A)$.

The simplest properties of polars are as follows:

- (1) If $A \subset B \subset X$ then $\pi(A) \supset \pi(B)$;
- (2) For every $A \subset X$ the inclusions hold:

$$A \subset \pi^{-1}(\pi(A)); \quad A \times \pi(A) \subset F;$$

- (3) If $A \times B \subset F$ then $B \subset \pi(A)$ and $A \subset \pi^{-1}(B)$;
- (4) If $(A_\xi)_{\xi \in \Xi}$ is a nonempty family of subsets of X then $\pi(\bigcup_{\xi \in \Xi} A_\xi) = \bigcap_{\xi \in \Xi} \pi(A_\xi)$;
- (5) If $A \subset X$ and $B \subset Y$ then $\pi(A) = \pi(\pi^{-1}(\pi(A)))$ and $\pi^{-1}(B) = \pi^{-1}(\pi(\pi^{-1}(B)))$.

3.1.10. Provided that $\text{Rel}(X) \wedge ((X \cap Y^2) \circ (X \cap Y^2) \subset X)$, we call X a *transitive* relation on Y . A relation X is *reflexive* (over Y) if $\text{Rel}(X) \wedge (I_Y \subset X)$. A relation X is *symmetric* if $X = X^{-1}$. Finally, we say that “ X is an *antisymmetric* relation on Y ” if $\text{Rel}(X) \wedge ((X \cap X^{-1}) \cap Y^2 \subset I_Y)$. As usual, we use the conventional abbreviation $Y^2 := Y \times Y$.

A reflexive and transitive relation on Y is a *preorder* on Y . An antisymmetric preorder on Y is an *order* or *ordering* on Y . A symmetric preorder is an *equivalence*.

Other terms are also applied that are now in common parlance. Recall in particular that an order X on Y is *total* or *linear*, while Y itself is called a *chain* (relative to X), whenever $Y^2 \subset X \cup X^{-1}$. If each nonempty subset of the set Y has a least element (relative to the order of X) then we say that X *well orders* Y . The terms *well-ordered* and *well-orderable* are understood correspondingly.

3.1.11. Quantifiers are *bounded* if they appear in the text as $(\forall x \in y)$ or $(\exists x \in y)$. The formulas of set theory (and, generally speaking, of every first-order theory) are classified according to how they use bounded and unbounded quantifiers.

Of especial importance to our exposition are the class of bounded formulas or Σ_0 -*formulas* and the class of the so-called Σ_1 -*formulas*. Recall that a formula φ is *bounded* provided that each quantifier in φ is bounded. Say that φ is of *class* Σ_1 or a Σ_1 -*formula* if φ results from atomic formulas and their negations by using only the logical operations $\wedge, \vee, (\forall x \in y)$, and $(\exists x)$.

Clearly, every bounded formula is of class Σ_1 . However, it is false that every Σ_1 -formula is bounded. Moreover, there are formulas not belonging to the class Σ_1 . The corresponding examples follow. We start with bounded formulas.

3.1.12. The proposition $z = \{x, y\}$ amounts to the bounded formula

$$x \in z \wedge y \in z \wedge (\forall u \in z)(u = x \vee u = y).$$

So, the definition of ordered pair is a bounded formula. The same holds for the definition of product since we may rewrite $Z = X \times Y$ as

$$(\forall z \in Z)(\exists x \in X)(\exists y \in Y)(z = (x, y)) \wedge (\forall x \in X)(\forall y \in Y)(\exists z \in Z)(z = (x, y)).$$

Another bounded formula reads “a mapping F from X to Y ” (see 3.1.8). Indeed, the above shows that $F \subset X \times Y$ is a bounded formula. Moreover, bounded are the expressions $\text{dom}(F) = X$ and $\text{Un}(F)$, equivalent to the respective formulas

$$\begin{aligned} & (\forall x \in X)(\exists y \in Y)(\exists z \in F)(z = (x, y)); \\ & (\forall z_1 \in F)(\forall z_2 \in F)(\forall x \in X)(\forall y_1 \in Y)(\forall y_2 \in Y) \\ & \quad (z_1 = (x, y_1) \wedge z_2 = (x, y_2) \rightarrow y_1 = y_2). \end{aligned}$$

3.1.13. The statements, that x and y are *equipollent*, or *equipotent*, or x and y have the same *cardinality*, symbolically,

$$x \simeq y,$$

each implying that there is a bijection between x and y , are all equivalent to the following Σ_1 -formula:

$$(\exists f)(f : x \rightarrow y \wedge \text{im}(f) = y \wedge \text{Un}(f^{-1})).$$

However, this fact is not expressible by a bounded formula. The notion of abstract relation gives another Σ_1 -formula:

$$\text{Rel}(X) := (\forall u \in X)(\exists v)(\exists w)(u = (v, w)).$$

Beyond the class Σ_1 lies the following formula stating that a set y is equipollent to none of its members:

$$(\forall x \in y) \neg(x \simeq y).$$

3.1.14. COMMENTS.

(1) It goes without saying that we may vary not only the special axioms of a first-order theory (see 3.1.4) but also its logical part, i.e., the logical axioms and rules of inference. The collections of the so-resulting theorems may essentially differ from each other. For instance, eliminating the law of the excluded middle from the axioms of propositional calculus, we arrive at intuitionistic propositional calculus. Intuitionistic predicate calculus (see [133, 201]) appears in a similar way.

(2) The modern formal logic was grown in the course of the evolution of philosophical and mathematical thought with immense difficulties. The classical predicate calculus originates with the Aristotle syllogistic whereas the origin of intuitionistic logic belongs elsewhere. Other logical systems, different essentially from the two systems, were invented in various times for various purposes. For instance, an ancient Indian logic had three types of negation, expressing the ideas: something has never exist and cannot happen now, something was available for the time being but is absent now, and something happens now but will disappear soon.

(3) As is seen from 3.1.6 and 3.1.7, abbreviations may appear in formulas, in other abbreviations, in abbreviations of abbreviations, etc. Invention of abbreviating symbols is an art in its own right, and as such it can never be formalized completely. Nevertheless, systemizing and codifying the rules for abbreviation is at the request of both theory and practice. Some advice (on exact descriptions, introduction of function letters, etc.) is available in the literature [58, 175, 238].

3.2. Zermelo–Fraenkel Set Theory

As has been noted in 3.1.4, the axioms of set theory include the general logical axioms of predicate calculus which postulate the classical rules for logical inference. Below we list the special axioms of set theory, ZF_1 – ZF_6 and AC. The theory proclaiming ZF_1 – ZF_6 as special axioms is called *Zermelo–Fraenkel set theory* and denoted by ZF. Enriching ZF with the axiom of choice AC, we come to a wider theory denoted by ZFC and still called Zermelo–Fraenkel set theory (the words “with choice” are added rarely). Note that we supply the formal axioms below with their verbal statements in the wake of the cantorians views of sets.

3.2.1. We often encounter the terms “*property*” and “*class*” dealing with ZFC. We now elucidate their formal statuses. Consider a formula $\varphi = \varphi(x)$ of ZFC (in symbols, $\varphi \in (\text{ZFC})$). Instead of the text $\varphi(y)$ we write $y \in \{x : \varphi(x)\}$. In other words, we use the so-called *Church schema* for classification:

$$y \in \{x : \varphi(x)\} := \varphi(y).$$

The expression $y \in \{x : \varphi(x)\}$ means in the language of ZFC that y has the property φ or, in other words, y belongs to the class $\{x : \varphi(x)\}$. Bearing this in mind, we say that a property, a formula, and a class mean the same in ZFC. We has already applied the Church schema in 3.1.6 and 3.1.7. Working within ZFC, we conveniently use many current abbreviations:

$$\begin{aligned} \mathbf{U} &:= \{x : x = x\} \text{ is the } \textit{universe of discourse} \text{ or the } \textit{class of all sets}; \\ \{x : \varphi(x)\} \in \mathbf{U} &:= (\exists z)(\forall y)\varphi(y) \leftrightarrow y \in z; \\ \{x : \varphi(x), \psi(x)\} &:= \{x : \varphi(x)\} \cap \{x : \psi(x)\}; \\ x \cup y &:= \bigcup \{x, y\}, \quad x \cap y \cap z := \bigcap \{x, y, z\} \dots \end{aligned}$$

We are now ready to formulate the special axioms of ZFC.

3.2.2. Axiom of Extensionality ZF_1 . *Two sets are equal if and only if they contain the same elements:*

$$(\forall x)(\forall y)(\forall z)((z \in x \leftrightarrow z \in y) \leftrightarrow x = y).$$

Note that we may replace the last equivalence by \rightarrow without loss of scope, since the reverse implication is a theorem of predicate calculus.

3.2.3. Axiom of Union ZF_2 . *The union of a set of sets is also a set:*

$$(\forall x)(\exists y)(\forall z)(\exists u)((u \in z \wedge z \in x) \leftrightarrow z \in y).$$

With the abbreviations of 3.1.6 and 3.2.1, ZF_2 takes the form

$$(\forall x) \bigcup x \in \mathbf{U}.$$

3.2.4. Axiom of Powerset ZF_3 . All subsets of each set comprise a new set:

$$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow (\forall u)(u \in z \rightarrow u \in x)).$$

In short,

$$(\forall x) \mathcal{P}(x) \in \mathbf{U}.$$

This axiom is also referred to as the *axiom of powers*.

3.2.5. Axiom of Replacement ZF_4^φ . The image of a set under each bijective mapping is a set again:

$$\begin{aligned} & (\forall x)(\forall y)(\forall z)(\varphi(x, y) \wedge \varphi(x, z) \rightarrow y = z) \\ & \rightarrow (\forall a)(\exists b)((\exists s \in a)(\exists t) \varphi(s, t) \leftrightarrow t \in b). \end{aligned}$$

In short,

$$\begin{aligned} & (\forall x)(\forall y)(\forall z)(\varphi(x, y) \wedge \varphi(x, z) \rightarrow y = z) \\ & \rightarrow (\forall a)(\{v : (\exists u \in a)\varphi(u, v)\} \in \mathbf{U}). \end{aligned}$$

Here φ is a formula of ZFC containing no free occurrences of a . Note that ZF_4^φ is a schema for infinitely many axioms since a separate axiom appears with an arbitrary choice of $\varphi \in (\text{ZFC})$. Bearing in mind this peculiarity, we often abstain from using a more precise term “axiom-schema” and continue speaking about the axiom of replacement for the sake of brevity and uniformity.

Note a few useful corollaries of ZF_4^φ .

3.2.6. Let $\psi = \psi(z)$ be a formula of ZFC. Given a set x , we may arrange a subset of x by collecting the members of x with the property ψ , namely,

$$(\forall x) \{z \in x : \psi(z)\} \in \mathbf{U}.$$

Our claim is ZF_4^φ , with $\psi(u) \wedge (u = v)$ playing the role of φ . This particular form of the axiom of replacement is often called the *axiom of separation* or *comprehension*.

3.2.7. Applying ZF_4^φ to the formula

$$\varphi(u, v) := (u = \emptyset \rightarrow v = x) \wedge (u \neq \emptyset \rightarrow v = y)$$

and the set $z := \mathcal{P}(\mathcal{P}(\emptyset))$, we deduce that the unordered pair $\{x, y\}$ of two sets (cf. 3.1.7) is also a set. This assertion is often referred to as the *axiom of pairing*.

3.2.8. Axiom of Infinity \mathbf{ZF}_5 . *There is at least one infinite set:*

$$(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \rightarrow y \cup \{y\} \in x)).$$

In other words, there is a set x such that $\emptyset \in x$, $\{\emptyset\} \in x$, $\{\emptyset, \{\emptyset\}\} \in x$, $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\} \in x$, etc. The cute reader will observe a tiny gap between formal and informal statements of the axiom of infinity. The vigilant reader might suspect the abuse of the term “infinity.” In fact, the axiom of infinity belongs to the basic cantorinan doctrines and so some mystery is inevitable and welcome in this respect.

3.2.9. Axiom of Regularity \mathbf{ZF}_6 . *Each nonempty set is disjoint from at least one of its members:*

$$(\forall x)(x \neq \emptyset \rightarrow (\exists y)(y \in x \wedge y \cap x = \emptyset)).$$

Another name for the axiom of regularity is the *axiom of foundation*.

Applying \mathbf{ZF}_6 to a *singleton*, i.e., a one-point set $x := \{y\}$, we see that $y \notin y$. Speaking a bit prematurely, we may note, on taking $x := \{x_1, \dots, x_n\}$, that there are no infinitely decreasing \in -sequences $x_1 \ni x_2 \ni \dots \ni x_n \ni \dots$.

3.2.10. Axiom of Choice AC. *To each set x there is a choice function on x ; i.e., a single-valued correspondence assigning an element of X to each nonempty member of X ; i.e.,*

$$(\forall x)(\exists f)(\text{Fnc}(f) \wedge x \subset \text{dom}(f)) \wedge (\forall y \in x) y \neq \emptyset \rightarrow f(y) \in y.$$

Set theory has many propositions equivalent to AC (cf. [196]). We recall the two most popular among them.

Zermelo Theorem (the well-ordering principle). *Every set is well-orderable.*

Kuratowski–Zorn Lemma (the maximality principle). *Let M be a (partially) ordered set whose every chain has an upper bound. Then to each $x \in M$ there is a maximal element $m \in M$ satisfying $m \geq x$.*

3.2.11. The axiomatics of ZFC enables us to find a concrete presentation for the class of all sets in the shape of the “von Neumann universe.” We start with the empty set. Each step of the construction consists in uniting the powersets of all available sets, thus making the stage for the next step. The transfinite repetition of these steps yields the von Neumann universe. Classes (in a “Platonic” sense) are viewed as external objects lying beyond the universe of discourse. Pursuing this approach, we consider a class as a family of sets obeying some set-theoretic property that is expressed by a formula of Zermelo–Fraenkel set theory. Therefore, the class consisting of some members of a certain set is a set itself (by the axiom of replacement). A formally sound definition of the von Neumann universe requires preliminary acquaintance with the notions of ordinal and cumulative hierarchy. We now turn to a minimum of prerequisites to these objects.

3.2.12. A set x is *transitive* if each member of x is a subset of x . A set x is an *ordinal* if x is transitive and totally ordered by the membership relation \in . These definitions look in symbolic form as follows:

$$\begin{aligned}\text{Tr}(x) &:= (\forall y \in x)(y \subset x) := \text{“}x \text{ is a transitive set”}; \\ \text{Ord}(x) &:= \text{Tr}(x) \wedge (\forall y \in x)(\forall z \in x) \\ & \quad (y \in z \vee z \in y \vee z = y) := \text{“}x \text{ is an ordinal.”}\end{aligned}$$

We commonly denote ordinals by lower case Greek letters. Every ordinal is endowed with the natural order by membership: given $\beta, \gamma \in \alpha$, we put

$$\gamma \leq \beta \leftrightarrow \gamma \in \beta \vee \gamma = \beta.$$

The class of all ordinals is denoted by On . So, $\text{On} := \{\alpha : \text{Ord}(\alpha)\}$.

An ordinal is a *well-ordered set*; i.e., it is totally ordered and its every subset has the least element (which is ensured by the axiom of regularity). We can easily see that

$$\begin{aligned}\alpha \in \text{On} \wedge \beta \in \text{On} &\rightarrow \alpha \in \beta \vee \alpha = \beta \vee \beta \in \alpha; \\ \alpha \in \text{On} \wedge \beta \in \alpha &\rightarrow \beta \in \text{On}; \\ \alpha \in \text{On} &\rightarrow \alpha \cup \{\alpha\} \in \text{On}; \\ &\text{Ord}(\emptyset).\end{aligned}$$

The ordinal $\alpha + 1 := \alpha \cup \{\alpha\}$ is called the *successor* of α or the *son* of α . A nonzero ordinal other than a successor is a *limit* ordinal. The following notation is common:

$$\begin{aligned}K_{\text{I}} &:= \{\alpha \in \text{On} : (\exists \beta) \text{Ord}(\beta) \wedge \alpha = \beta + 1 \vee \alpha = \emptyset\}; \\ K_{\text{II}} &:= \{\alpha \in \text{On} : \alpha \text{ is a limit ordinal}\}; \\ 0 &:= \emptyset, \quad 1 := 0 + 1, \quad 2 := 1 + 1, \dots, \\ \omega &:= \{0, 1, 2, \dots\} = 0 \cup \mathbb{N}.\end{aligned}$$

This is a right place to recall that the *continuum* we talk about in this book from time to time is simply the powerset of ω .

3.2.13. It is worth observing that ZFC enables us to prove the properties of ordinals well known at a naive level. In particular, ZFC legitimizes transfinite induction and recursion. We now define the von Neumann universe, purposefully omitting formalities.

Given an ordinal α , put

$$V_\alpha := \bigcup_{\beta < \alpha} \mathcal{P}(V_\beta),$$

i.e., $V_\alpha = \{x : (\exists \beta)(\beta \in \alpha \wedge x \subset V_\beta)\}$. More explicitly,

$$\begin{aligned} V_0 &:= \emptyset; \\ V_{\alpha+1} &:= \mathcal{P}(V_\alpha); \\ V_\beta &:= \bigcup_{\alpha < \beta} V_\alpha \text{ if } \beta \in K_{\text{II}}. \end{aligned}$$

Assign

$$\mathbb{V} := \bigcup_{\alpha \in \text{On}} V_\alpha.$$

Of principal importance is the following theorem, ensuing from the axiom of regularity:

$$(\forall x)(\exists \alpha)(\text{Ord}(\alpha) \wedge x \in V_\alpha).$$

In shorter symbols,

$$\mathbf{U} = \mathbb{V}.$$

Alternatively, we express this fact as follows: “*The class of all sets is the von Neumann universe,*” or “*every set is well-founded.*”

The von Neumann universe \mathbb{V} , also called the *sets*, is customarily viewed as a pyramid “upside down,” that is, a pyramid standing on its vertex which is the empty set (Fig. 3). It is helpful to look at a few “lower floors” of the von Neumann universe:

$$\begin{aligned} V_0 &= \emptyset, \quad V_1 = \{\emptyset\}, \quad V_2 = \{\emptyset, \{\emptyset\}\}, \dots, \\ V_\omega &= \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots\}, \dots \end{aligned}$$

The representation of the von Neumann universe \mathbb{V} as the “*cumulative hierarchy*” of $(V_\alpha)_{\alpha \in \text{On}}$ makes it possible to introduce the concept of the *ordinal rank* or simply the *rank* of a set. Namely, given a set x , put

$$\text{rank}(x) := \text{a least ordinal } \alpha \text{ such that } x \in V_{\alpha+1}.$$

It is easy to prove that

$$\begin{aligned} a \in b &\rightarrow \text{rank}(a) < \text{rank}(b); \\ \text{Ord}(\alpha) &\rightarrow \text{rank}(\alpha) = \alpha; \\ (\forall x)(\forall y) \text{rank}(y) < \text{rank}(x) &\rightarrow (\varphi(y) \rightarrow \varphi(x)) \rightarrow (\forall x) \varphi(x), \end{aligned}$$

where φ is a formula of ZFC. The preceding theorem (or, more precisely, the schema of theorems) is called the *principle of induction on rank*.

3.2.14. Two sets are *equipollent*, or *equipotent*, or *of the same cardinality* if there is a bijection of one of them onto the other. An ordinal that is equipotent to no preceding ordinal is a *cardinal*. Each natural is a cardinal.

A cardinal other than a natural is an *infinite* cardinal. Therefore, ω is the least infinite cardinal.

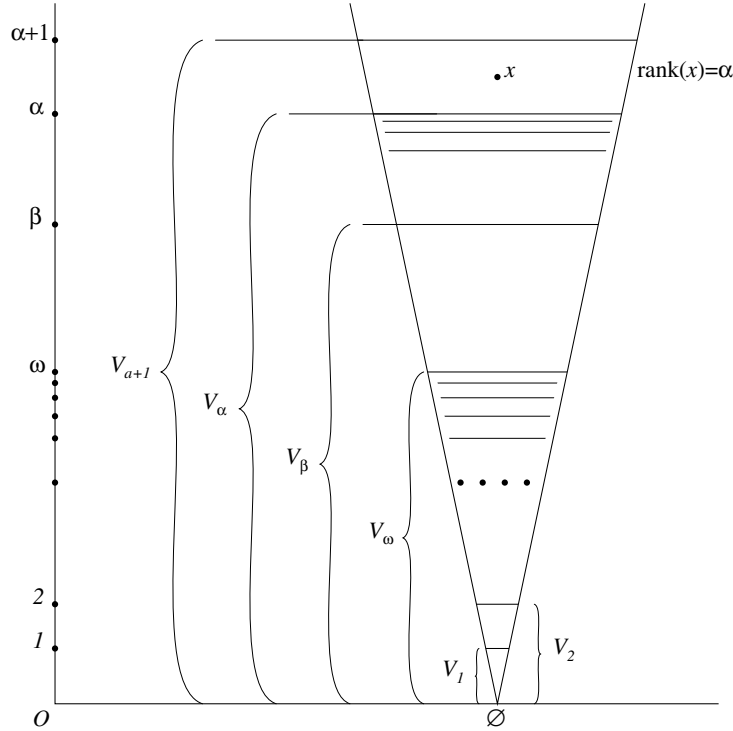


Fig. 3

The *Hartogs number* $\mathcal{H}(x)$ of a set x is the least of the ordinals α such that there is no injection from α to x . Clearly, $\mathcal{H}(x)$ is a cardinal for every x . Moreover, the Hartogs number of each ordinal α is the least of the cardinals strictly greater than α .

By recursion we define the *alephic scale*:

$$\aleph_0 := \omega_0 = \omega;$$

$$\aleph_{\alpha+1} := \omega_{\alpha+1} = \mathcal{H}(\omega_\alpha);$$

$$\aleph_\beta := \omega_\beta := \sup\{\omega_\alpha : \alpha < \beta\} \text{ if } \beta \in K_{II}.$$

The following hold:

- (1) *Infinite cardinals form a well-ordered proper class;*

- (2) The mapping $\alpha \mapsto \omega_\alpha$ is an order isomorphism between the class of ordinals and the class of infinite cardinals;
- (3) There is a mapping $|\cdot|$ from the universal class \mathbf{U} onto the class of all cardinals such that the sets x and $|x|$ are equipollent for all $x \in \mathbf{U}$.

The cardinal $|x|$ is called the *cardinality* or the *cardinal number* of a set x . Hence, each set is equipollent to a unique cardinal which is its cardinality.

A set x is *countable* provided that $|x| = \omega_0 := \omega$, and x is *at most countable* provided that $|x| \leq \omega_0$.

Given an ordinal α , we denote by $\mathbf{2}^{\omega_\alpha}$ the cardinality of $\mathcal{P}(\omega_\alpha)$; i.e., $\mathbf{2}^{\omega_\alpha} := |\mathcal{P}(\omega_\alpha)|$. This denotation is justified by the fact that $\mathbf{2}^x$ and $\mathcal{P}(x)$ are equipollent for all x , with $\mathbf{2}^x$ standing for the class of all mappings from x to $\mathbf{2}$.

A theorem, proven by Cantor, states that $|x| < |\mathbf{2}^x|$ for whatever set x . In particular, $\omega_\alpha < \mathbf{2}^{\omega_\alpha}$ for each ordinal α . Moreover, $\omega_{\alpha+1} \leq \mathbf{2}^{\omega_\alpha}$.

The *generalized problem of the continuum* asks whether or not there are intermediate cardinals between $\omega_{\alpha+1}$ and $\mathbf{2}^{\omega_\alpha}$; i.e., whether or not the equality $\omega_{\alpha+1} = \mathbf{2}^{\omega_\alpha}$ holds. For $\alpha = 0$ this is the classical *continuum problem*.

The *continuum hypothesis* or CH is the equality $\omega_1 = \mathbf{2}^\omega$. Similarly, the *generalized continuum hypothesis* or GCH is the equality $\omega_{\alpha+1} = \mathbf{2}^{\omega_\alpha}$ for all $\alpha \in \text{On}$.

3.2.15. In the sequel we will make use of the following technical result of profound importance which is often called the *reflection principle*. In a sense, this result shows that all “particular” set-theoretic events happen to sets of bounded rank.

Montague–Levy Theorem. Let $\varphi := \varphi(x_1, \dots, x_n)$ be a formula of ZFC. Take an ordinal α . Then there is an ordinal β such that $\beta > \alpha$ and

$$(\forall x_1, \dots, x_n \in V^\beta) \varphi(x_1, \dots, x_n) \leftrightarrow \varphi^{V^\beta}(x_1, \dots, x_n),$$

where φ^{V^β} is the relativization of φ to V^β .

◁ Assume that the prenex normal form of φ looks as follows

$$\varphi = (Q_1 y_1) \dots (Q_m y_m) \psi(x_1, \dots, x_n, y_1, \dots, y_m).$$

In other words, ψ is quantifier-free and $Q_k \in \{\exists, \forall\}$.

Put

$$\psi_k := (Q_{k+1} y_{k+1}) \dots (Q_m y_m) \psi$$

for $k := 0, \dots, m$. With due precaution, it is possible to conclude that

$$\psi_k = \psi_k(x_1, \dots, x_n, y_1, \dots, y_{k-1}).$$

Assigning a collection of unbound variables in ψ_k , find the least ordinal γ satisfying

$$(\exists y_k)\psi_k \rightarrow (\exists y_k \in V^\gamma)\psi_k$$

provided that $Q_k = \exists$ and

$$(\exists y_k)\neg\psi_k \rightarrow (\exists y_k \in V^\gamma)\neg\psi_k$$

provided that $Q_k = \forall$. Put

$$g_k(x_1, \dots, x_n, y_1, \dots, y_{k-1}) := \gamma.$$

Given an ordinal α and using the axiom of replacement 3.2.5, find the set $A_k(\alpha)$ so that

$$\{g_k(x_1, \dots, x_n, y_1, \dots, y_{k-1}) : x_1, \dots, x_n \in V^\alpha; y_1, \dots, y_{k-1} \in V^\alpha\}.$$

Put

$$f_k(\alpha) := \sup\{\alpha + 1, (\sup A_k(\alpha)) + 1\}.$$

Using the so-constructed ordinal-valued functions, we successively put

$$\begin{aligned} f^{(0)}(\alpha) &:= \alpha; \\ f^{(1)}(\alpha) &:= \sup\{f_1(\alpha), \dots, f_m(\alpha)\}; \\ f^{(s+1)}(\alpha) &:= f^{(1)}(f^{(s)}(\alpha)) \quad (s \in \mathbb{N}). \end{aligned}$$

And, finally,

$$f(\alpha) := \sup_{s \in \mathbb{N}} f^{(s)}(\alpha).$$

Clearly, $f(\alpha)$ is a limit ordinal greater than α for all α . Moreover,

$$g_k(x_1, \dots, x_n, y_1, \dots, y_{k-1}) < f(\alpha)$$

for all $x_1, \dots, x_n, y_1, \dots, y_m \in V^{f(\alpha)}$ and $1 \leq k \leq m$.

Putting $\beta := f(\alpha)$, considering that $\psi_{k-1} = (Q_k y_k)\psi_k$, and using the definition of g_k , proceed successively as follows:

$$\begin{aligned} &\psi_m = \psi_m^{V^\beta} \\ &\rightarrow (\psi_{m-1} \leftrightarrow (Q_m y_m \in V^\beta)\psi_m) \\ &\quad \rightarrow (\psi_{m-1} \leftrightarrow \psi_{m-1}^{V^\beta}) \\ &\quad \rightarrow \dots \rightarrow \psi_1 \leftrightarrow \psi_1^{V^\beta} \\ &\rightarrow (Q_1 y_1)\psi_1 \leftrightarrow (Q_1 y_1 \in V^\beta)\psi_1 \\ &\quad \rightarrow \psi_0 \leftrightarrow \psi_0^{V^\beta} \\ &\rightarrow \varphi(x_1, \dots, x_n) \leftrightarrow \varphi^{V^\beta}(x_1, \dots, x_n). \end{aligned}$$

This ends the proof. \triangleright

Corollary. Let $\varphi_1, \dots, \varphi_m$ be formulas of ZFC whose unbound variables are listed among x_1, \dots, x_n . Take $\alpha \in \text{On}$. Then

$$(\exists \beta > \alpha)(\forall x_1, \dots, x_n \in V^\beta)(\varphi_1 \leftrightarrow \varphi_1^{V^\beta}) \wedge \dots \wedge (\varphi_m \leftrightarrow \varphi_m^{V^\beta}).$$

◁ Putting

$$\varphi(t, x_1, \dots, x_n) = (t = 1 \wedge \varphi_1) \vee (t = 2 \wedge \varphi_2) \vee \dots \vee (t = m \wedge \varphi_m),$$

apply the above theorem. ▷

3.2.16. Study of various models of set theory often involves the *ultrapower construction*. We now provide some details that will be needed to the reader who intends to elaborate formalities of the status of nonstandard set theories.

Assume that U is some set and ε is some relation on U . In the context of set theory, such a couple (U, ε) is often referred to as a *universet* or *universoid*. In this event, we will sometimes write $x\varepsilon y$ instead of $(x, y) \in \varepsilon$.

Let $\varphi = \varphi(x_1, \dots, x_n)$ be a formula of ZFC. Interpret ε as membership and restrict all quantifiers of φ to U . Assuming that $\varphi(x_1, \dots, x_n)$ for $x_1, \dots, x_n \in U$, we will write $(U, \varepsilon) \models \varphi(x_1, \dots, x_n)$ or $\varphi^{(U, \varepsilon)}(x_1, \dots, x_n)$ or even φ^U and speak of the *relativization* of φ to U . Other abbreviations are also popular.

Consider the power $\mathfrak{X} := X^\mathcal{E}$ of some set X , where \mathcal{E} is some index set (for the sake of convenience, we assume X and \mathcal{E} nonempty). Given $x_1, \dots, x_n \in \mathfrak{X}$ and $\varphi = \varphi(x_1, \dots, x_n) \in (\text{ZFC})$, put

$$\llbracket \varphi(x_1, \dots, x_n) \rrbracket = \{e \in \mathcal{E} : \varphi^X(x_1(e), \dots, x_n(e))\},$$

where φ^X is the relativization of φ to X .

Assume further that \mathcal{F} is a filter on \mathcal{E} and

$$f \sim_{\mathcal{F}} g := \llbracket f = g \rrbracket \in \mathcal{F} \quad (f, g \in \mathfrak{X}).$$

Denote the quotient set $\mathfrak{X} / \sim_{\mathcal{F}}$ by ${}^{\mathcal{F}}X$ and let ${}^{\mathcal{F}}f$ stand for the coset of f . Clearly,

$$\llbracket f' \varepsilon g' \rrbracket = \llbracket f = f' \rrbracket \cap \llbracket g = g' \rrbracket \cap \llbracket f' \varepsilon g' \rrbracket$$

for $f \sim_{\mathcal{F}} f'$ and $g \sim_{\mathcal{F}} g'$.

Therefore, $\llbracket f \varepsilon g \rrbracket \in \mathcal{F} \leftrightarrow \llbracket f' \varepsilon g' \rrbracket \in \mathcal{F}$. In other words, we have soundly defined on ${}^{\mathcal{F}}X$ the following relation

$$\mathcal{F}_\varepsilon := \{({}^{\mathcal{F}}f, {}^{\mathcal{F}}g) \in ({}^{\mathcal{F}}X)^2 : \llbracket f \varepsilon g \rrbracket \in \mathcal{F}\}.$$

It is easy to see that

$$\mathcal{F}_\varepsilon = \sim_{\mathcal{F}} \circ \varepsilon_{\mathfrak{X}} \circ \sim_{\mathcal{F}}$$

for some appropriate interpretation $\varepsilon_{\mathfrak{X}}$ on \mathfrak{X} of the membership relation. We call ${}^{\mathcal{F}}X$ the *filtered power* of X by \mathcal{F} . If \mathcal{F} is an ultrafilter then ${}^{\mathcal{F}}X$ is the *ultrapower* of X by \mathcal{F} .

Given $f, g \in \mathfrak{X}$, from these definitions we infer that

$$\begin{aligned} {}^{\mathcal{F}}f \mathcal{F}_\varepsilon {}^{\mathcal{F}}g &\leftrightarrow \llbracket f \varepsilon g \rrbracket \in \mathcal{F}; \\ {}^{\mathcal{F}}f = {}^{\mathcal{F}}g &\leftrightarrow \llbracket f = g \rrbracket \in \mathcal{F}. \end{aligned}$$

In other words,

$$({}^{\mathcal{F}}X, \mathcal{F}_\varepsilon) \models \varphi^{{}^{\mathcal{F}}X}({}^{\mathcal{F}}f, {}^{\mathcal{F}}g) \leftrightarrow \llbracket \varphi(f, g) \rrbracket \in \mathcal{F}$$

for every atomic formula $\varphi = \varphi(x, y)$ of ZFC.

Given $x \in X$, put $\bar{x}(e) := x$ for all $e \in \mathcal{E}$ and $*x := {}^{\mathcal{F}}x$. Note that $*x = *y \leftrightarrow x = y$ and $*x \mathcal{F}_\varepsilon *y \leftrightarrow x \varepsilon y$. It turns out that such effects are routine. To describe them, the following definition is in order.

Let $\varphi = \varphi(x_1, \dots, x_n)$ be an arbitrary formula of ZFC. We say that φ is *filtered* (with respect to X , \mathcal{E} , and \mathcal{F}) provided that

$$({}^{\mathcal{F}}X, \mathcal{F}_\varepsilon) \models \varphi^{{}^{\mathcal{F}}X}({}^{\mathcal{F}}f_1, \dots, {}^{\mathcal{F}}f_n) \leftrightarrow \llbracket \varphi(f_1, \dots, f_n) \rrbracket \in \mathcal{F}$$

for all $f_1, \dots, f_n \in \mathfrak{X}$.

Loš Theorem. *Every formula of ZFC is filtered with respect to an arbitrary ultrafilter.*

◁ Since all atomic formulas are filtered, it suffices to check that the application of propositional connectives and quantification preserve filteredness. If φ is a filtered formula then $\neg\varphi$ is filtered by the “golden” property of every ultrafilter: $F \in \mathcal{F} \leftrightarrow F' := \mathcal{E} - F \notin \mathcal{F}$. We will thus establish the following (absolutely unavoidable) fact: If $\psi(y) := (\forall x)\varphi(x, y)$ and φ is filtered then so is ψ .

To this end, assume that $\llbracket \psi(y) \rrbracket \in \mathcal{F}$ for $y \in \mathcal{F}$ and $x \in {}^{\mathcal{F}}X$. Clearly, $\llbracket \psi(y) \rrbracket \subset \llbracket \varphi(x, y) \rrbracket$ and so, $({}^{\mathcal{F}}X, \mathcal{F}_\varepsilon) \models {}^{\mathcal{F}}\varphi(x, y)$. Since x is arbitrary, therefore, $({}^{\mathcal{F}}X, \mathcal{F}_\varepsilon) \models (\forall x) \varphi^{\mathcal{F}}(x, y)$.

Assume finally that $x, y \in {}^{\mathcal{F}}X$ implies $\varphi^{{}^{\mathcal{F}}X}(x, y)$, i.e. $\llbracket \varphi(x, y) \rrbracket \in \mathcal{F}$. Check that $B := \llbracket (\forall x)\varphi(x, y) \rrbracket$ belongs to \mathcal{F} as well. Indeed, to $e \in \mathcal{E} - B := B'$ there is some $\bar{x}(e)$ satisfying $\neg\varphi(x(e), y(e))$. Take an arbitrary x_0 in \mathfrak{X} . Put $\bar{\bar{x}}(e) := \bar{x}(e)$ whenever $e \in B'$ and $\bar{\bar{x}}(e) := x_0(e)$ otherwise. Obviously, $\llbracket \varphi(\bar{\bar{x}}, y) \rrbracket \subset \mathcal{E} - B' = B$. Since $\llbracket \varphi(\bar{\bar{x}}, y) \rrbracket \in \mathcal{F}$, it follows that $B \in \mathcal{F}$. The proof is complete. ▷

Corollary 1. *Let X be a nonempty set and let $*X$ be some ultrapower of X . Then*

$$\varphi^X(x_1, \dots, x_n) \leftrightarrow \varphi^{*X}(*x_1, \dots, *x_n)$$

for $x_1, \dots, x_n \in X$ and $\varphi \in (\text{ZFC})$.

◁ By the Łoś Theorem $\varphi^{*X}(*x_1, \dots, *x_n) \leftrightarrow \llbracket \varphi(\bar{x}_1, \dots, \bar{x}_n) \rrbracket \in \mathcal{F}$, where \mathcal{F} is the ultrafilter in question and $\bar{x}_k(e) = (x_k)$ for all $e \in \mathcal{E}$. From the definition of $\llbracket \cdot \rrbracket$ it follows that $\llbracket \varphi(\bar{x}_1, \dots, \bar{x}_n) \rrbracket \in \mathcal{F} \leftrightarrow \varphi^X(x_1, \dots, x_n)$, which completes the proof. ▷

Let X be an infinite set and let \mathcal{E} stand for the *cofinite filter* on X which consists of the complements to X of finite subsets of X . Assume further that \mathcal{F} is some ultrafilter finer than \mathcal{E} . The ultrapower $\mathcal{F}X$ is a *canonical enlargement* of X , denoted still by $*X$.

Corollary 2 (the weak idealization principle). *Let $\varphi = \varphi(x, y, x_1, \dots, x_n)$ be a formula of ZFC. Assume given elements $x_1, \dots, x_n \in X$ and a canonical enlargement $*X$ of X . Then*

$$\begin{aligned} & (\forall^{\text{fin}} A \subset X)(\exists b \in X)(\forall a \in A) \varphi^X(a, b, x_1, \dots, x_n) \\ & \rightarrow (\exists b \in *X)(\forall a \in A) \varphi^{*X}(*a, b, *x_1, \dots, *x_n). \end{aligned}$$

◁ To $e \in \mathcal{E}$ there is some $b(e)$ in X satisfying $(\forall a \in e) \varphi^X(a, b(e), x_1, \dots, x_n)$. In other words, using $b \in X^{\mathcal{E}}$, we see that

$$\llbracket \varphi(\bar{x}, b, \bar{x}_1, \dots, \bar{x}_n) \rrbracket \subset \{e \in \mathcal{E} : a \in e\}$$

where, as usual, $\bar{y}(e) := y$ for $y \in X$ and $e \in \mathcal{E}$.

By the Łoś Theorem, $\varphi^{*X}(*a, \mathcal{F}b, *x_1, \dots, *x_n)$. This ends the proof. ▷

Let Z be a nonempty subset of \tilde{Z} . This Z is a *Zermelo subset* of \tilde{Z} provided that

- (a) Z is transitive in \bar{Z} (i.e., $a \in \tilde{Z} \wedge b \in Z \wedge a \in b \rightarrow a \in Z$);
- (b) Z is closed under unordered pairing;
- (c) $a \in Z \rightarrow \bigcup a \in Z \wedge \mathcal{P}(a) \in Z$.

Let $(\tilde{Z}, \tilde{\varepsilon})$ be a universet. Assume given another universet (Z, ε) such that Z is a nonempty subset of \tilde{Z} and ε is the restriction of $\tilde{\varepsilon}$ to E^2 . In this event (Z, ε) is a *subuniverset* of $(\tilde{Z}, \tilde{\varepsilon})$.

Assume that Z models a Zermelo subset of \tilde{Z} if $\tilde{\varepsilon}$ is interpreted as membership. In this event Z is a *Zermelo universet* (in $(\tilde{Z}, \tilde{\varepsilon})$). The indication of \tilde{Z} is often omitted when this leads to no confusion.

Corollary 3. *Let $(\mathfrak{X}, \varepsilon)$ be a Zermelo universet and let $*\mathfrak{X}$ be some ultrapower of X . Assume further that $X \in \mathfrak{X}$, $Y \in *\mathfrak{X}$, and $\tilde{f} : X \rightarrow \tilde{Y}$ (i.e., \tilde{f} is an external function), with $\tilde{Y} := \{y : y^{*\mathfrak{X}} \in Y\}$. Then there is some f in $*\mathfrak{X}$ such that f is a function from $*X$ to Y inside $(*\mathfrak{X})$ and, moreover, $\tilde{f}(x) = f(*x)$ for all $x \in X$.*

◁ In case $\tilde{f} = \emptyset$, put $f := \emptyset$. If $\tilde{f} \neq \emptyset$ then $Y \neq \emptyset$. Assume that $Y = \mathcal{F}\mathcal{Y}_0$, where \mathcal{F} is the ultrafilter in question on an appropriate direction \mathcal{E} . In this event $\llbracket \mathcal{Y}_0 \neq \emptyset \rrbracket = \{e \in \mathcal{E} : \mathcal{Y}_0 \neq \emptyset\} \in \mathcal{F}$. Redefining $\mathcal{Y}_0(e)$, if need be, for $e \notin \llbracket \mathcal{Y}_0 \neq \emptyset \rrbracket$, we may assume that $Y = \mathcal{F}\mathcal{Y}$ and $\mathcal{Y}(e) \neq \emptyset$ for all $e \in \mathcal{E}$.

Suppose that $y \in \tilde{Y}$ and $y = \mathcal{F}\eta$. Clearly, $\llbracket \eta \in \mathcal{Y} \rrbracket \in \mathcal{F}$. Put $h(y)(e) := \eta(e)$ for $e \in \llbracket \eta \in \mathcal{Y} \rrbracket$ and define $h(y)$ at other values of e as, for instance, some member of $\mathcal{Y}(e) \in X$. What matters is the equality $\mathcal{F}h(y) = y$ holding irrespectively of this choice. Given $e \in \mathcal{E}$, define the function $g(e) : X \rightarrow \mathcal{Y}(e)$ by the rule $g(e)(x) := h(f(x))(e)$ where $x \in X$. The set $g(e) := \{(x, g(e)(x)) : x \in X\}$ is a member of \mathfrak{X} (since \mathfrak{X} is a Zermelo universet). We thus arrive at the element g of $\mathfrak{X}^{\mathcal{E}}$ satisfying $g : e \in \mathcal{E} \mapsto g(e) \in \mathfrak{X}$. It is evident also that $\llbracket g : \overline{X} \rightarrow \mathcal{Y} \rrbracket = \mathcal{E}$. Thus, $f := \mathcal{F}g$ is a function from $*X$ to Y by the Loš Theorem. To prove, we inspect the above arguments and note that

$$\begin{aligned} \tilde{f}(x) = f(*x) &\leftrightarrow f(*x) = \mathcal{F}h(\tilde{f}(x)) \\ &\leftrightarrow \llbracket g(\overline{x}) = h(\tilde{f}(x)) \rrbracket \in \mathcal{F} \end{aligned}$$

for $x \in X$.

Moreover, by definition

$$g(\overline{x})(e) = g(e)(x) = h(\tilde{f}(x))(e)$$

for all $e \in \mathcal{E}$, which finishes the proof. ▸

3.2.17. To study deeper properties we need a more abstract procedure which is known as the *ultralimit construction*. We give only a necessary minimum of information on ultralimits.

Assume that (U, ε) is a universet and $V := \mathcal{P}(U)$. Proceed by putting

$$\begin{aligned} f_U(u) &:= f(u) = \{v \in U : (v, u) \in \varepsilon\} \quad (u \in U); \\ E &:= \{(A, B) \in V \times V : (\exists a \in U)(A = f(a) \wedge a \in B)\}. \end{aligned}$$

Note that, given $v \subset U$, we infer by definition that

$$A \in f(v) \leftrightarrow (\exists u \in v)(A = f(u)) \leftrightarrow (A, v) \in E.$$

In other words,

$$f(v) = \{A \in V : (A, v) \in E\}.$$

The universet (V, E) is the *protoextension* of (U, ε) .

3.2.18. Let (U, ε) be a universet satisfying the axiom of extensionality and let (V, E) be the protoextension of (U, ε) . Then

- (1) (V, E) enjoys the axiom of extensionality;
- (2) $f := f_U : U \rightarrow U$ is an injective mapping and

$$(f(u), f(v)) \in E \leftrightarrow (u, v) \in \varepsilon.$$

- (3) $(\forall u \in U)(A, f(u)) \in E \leftrightarrow (\exists a)((a, u) \in \varepsilon \wedge A = f(a))$ for all $A \in V$.

\triangleleft (1): To check the axiom of extensionality in (V, E) , take $x, y \in V$ so that $(\forall z \in V)(z, x) \in E \rightarrow (z, y) \in E$. We have to validate that $x \subset y$. Take $w \in x, w \in U$. Then $(f(w), x) \in E$ and so $f(w) = f(\bar{w})$ and $\bar{w} \in y$ for some $\bar{w} \in U$. However, we already know that $\bar{w} = w$. Hence, $(\forall w \in U)w \in x \rightarrow w \in y$. We are done on recalling that $x, y \in \mathcal{P}(U)$.

(2): Note that $(u, v) \in \varepsilon$ amounts to $u \in f(v)$. Hence, on using the validity of the axiom of extensionality for ε in U , we infer that

$$\begin{aligned} f(u) = f(v) &\rightarrow ((\forall z \in U) z \in f(u) \leftrightarrow z \in f(v)) \\ &\rightarrow ((\forall z \in U)(z, u) \in \varepsilon \leftrightarrow (z, v) \in \varepsilon) \rightarrow u = v. \end{aligned}$$

Assuming now that $(f(u), f(v)) \in E$, by definition we see that $f(a) = f(a)$ and $a \in f(v)$ for some $a \in \bar{U}$. Since $a = u$ as proven; therefore, $(u, v) \in \varepsilon$. In turn, the implication $(u, v) \in \varepsilon \rightarrow (f(u), f(v)) \in E$ is clear (and does not imply extensionality in U) for we may take U as the element n that is requested by the definition of E .

(3): If $(a, u) \in \varepsilon$ and $A = f(a)$ then $a \in f(u)$. Hence, $(A, f(u)) \in E$ by definition. Conversely, from (2) we infer that

$$\begin{aligned} (A, f(u)) \in E &\rightarrow (\exists a \in U)A = f(a) \wedge a \in f(u) \\ &\rightarrow A = f(a) \wedge (f(a), f(u)) \in E \rightarrow (a, u) \in \varepsilon \wedge A = f(a), \end{aligned}$$

which ends the proof. \triangleright

Let (U, ε) be a universet satisfying the axiom of extensionality. Put $U_0 := U$ and $\varepsilon_0 := \varepsilon$. Using the above consecutively and granted (U_k, ε_k) , assign

$$\begin{aligned} U_{k+1} &:= \mathcal{P}(U_k); \\ f_k(u) &:= \{\bar{u} \in U_k : (\bar{u}, U) \in \varepsilon_k\} \quad (u \in U_k); \\ \varepsilon_{k+1} &:= \{(u, v) \in U_{k+1} \times U_{k+1} : (\exists a \in U_k)(u = f_k(a) \wedge a \in v)\}. \end{aligned}$$

We thus acquire the sequence of injections

$$U_0 \xrightarrow{f_1} U_1 \xrightarrow{f_2} U_2 \rightarrow \dots \rightarrow U_n \xrightarrow{f_{n+1}} U_{n+1} \rightarrow \dots$$

Hence,

$$\begin{aligned} g_n \circ \varepsilon_n \circ g_n^{-1} &= (g_{n+1} \circ f_{n+1}) \circ \varepsilon_n \circ (f_{n+1}^{-1} \circ g_{n+1}^{-1}) \\ &= g_{n+1} \circ (f_{n+1} \circ \varepsilon_n \circ f_{n+1}^{-1}) \circ g_{n+1}^{-1} \in g_{n+1} \circ \varepsilon_{n+1} \circ g_{n+1}^{-1}. \end{aligned}$$

We may thus assume that

$$(u, v) \in E \leftrightarrow (\exists n \in \mathbb{N})(u, v) \in g_{n+1} \circ \varepsilon_{n+1} \circ g_{n+1}^{-1}.$$

In this event, we also have

$$(u, v) \in g_{n+1} \circ \varepsilon_{n+1} \circ g_{n+1}^{-1}$$

for all $n \geq n_0$.

Clearly, $v = g_{n+1}(\bar{v})$ where $\bar{v} := g_{n+1}^{-1}(v) \in U_{n+1}$; and $u = g_{n+1}(\bar{u})$ where $\bar{u} := g_{n+1}^{-1}(u)$. Moreover, $\bar{u}\varepsilon_{n+1}\bar{v}$ and $\bar{v} \in U_{n+1}$, i.e. $\bar{v} = f_{n+1}(a)$ for some $a \in v_n$. Therefore, $(\bar{u}, f_{n+1}(a)) \in \varepsilon_{n+1}$. Consequently, from 3.2.18 (3) it follows that $\bar{u} = f_{n+1}(\bar{u})$ for some $\bar{u} \in U_n$. Hence, $u = g_{n+1}(\bar{u}) = g_{n+1}(f_{n+1}(\bar{u})) = g_n(\bar{u}) \in U'_n$. Since $(f_{n+1}(\bar{u}), \bar{v}) \in \varepsilon_{n+1}$; therefore, $\bar{u} \in \bar{v}$ by the definition of ε_{n+1} . We are done on observing that $g_n^{-1}(u) = \bar{u}$ and $g_{n+1}^{-1}(v) = \bar{v}$. \triangleright

3.2.20. *Let (V, E) be an external extension of a universet (U, ε) enjoying the axiom of extensionality. Then*

- (1) $(V, E) \models$ the axiom of extensionality;
- (2) $(V, E) \models$ the axiom of pairing;
- (3) $(V, E) \models$ the axiom of union;
- (4) $(V, E) \models$ the axiom of powerset;
- (5) $(V, E) \models$ the axiom-schema of comprehension;
- (6) $(V, E) \models$ the axiom of choice;
- (7) $(V, E) \models$ the axiom of the empty set;
- (8) $(V, E) \models$ the axiom of infinity;
- (9) $(\forall a, b \in U)((a, b) \in \varepsilon) \leftrightarrow (\iota(a), \iota(b)) \in E$;
- (10) $(\forall x, y \in V)((x, y) \in E \wedge y \in \iota(U) \rightarrow x \in \iota(U))$;
- (11) $(\forall \bar{U} \subset U)(\exists \bar{\bar{U}} \in V)(\forall v \in V)((v, \bar{\bar{U}}) \in E \leftrightarrow v \in \iota(\bar{\bar{U}}))$.

\triangleleft (1): From 3.2.18 (1) and the induction principle in (U_n, ε_n) we see the validity of the axiom of extensionality. It suffices to note the restriction of E to $U'_n \times U'_n$ coincides with ε'_n for all $n \in \mathbb{Z}_+$.

(2): Assume that $u, v \in U'_n$ with $u = g_n(x)$ and $v = g_n(y)$. The unordered $z := \{x, y\}$ is a member of U_{n+1} . Hence, $w := g_{n+1}(z)$ is a member of U'_{n+1} . Clearly, $(z, w) \in E \leftrightarrow z = u \vee z = v$.

(3): Take $u \in V$. We may assume that $u = g_{n+2}(x)$ and $x \in U_{n+2}$. Put

$$y := \bigcup \{f_{n+1}(z) : z \in f_{n+2}(x), z \in U_{n+1}\}.$$

Clearly, $y \in U_{n+1}$. Assigning $v := g_{n+1}(y)$, note that

$$(w, v) \in E \leftrightarrow (\exists \varkappa)(\varkappa, v) \in E \wedge (w, \varkappa) \in E$$

for $w \in V$.

Indeed, if $a \in U_{n+1}$ then

$$\begin{aligned} a \in y &\leftrightarrow ((\exists z \in U_{n+1})z \in f_{n+2}(x)) \wedge a \in f_{n+1}(z) \\ &\leftrightarrow (z, x) \in \varepsilon_{n+2} \wedge (a, z) \in \varepsilon_{n+1}. \end{aligned}$$

We are done on appealing to 3.2.19.

(4): Take $u \in V$ and let $u = g_n(x)$, with $x \in U_n$. Put $A := \{y \in U_n : f_{n+1}(y) \subset f_{n+1}(x)\}$. Check that the set $v := g_{n+2}(A)$ plays the role of the powerset of u inside (V, E) . To this end, observe first of all that

$$\begin{aligned} f_{n+1}(y) \subset f_{n+1}(x) &\leftrightarrow (\forall z \in V)(z, g_{n+1}(f_{n+1}(y))) \in E \rightarrow (z, u) \in E \\ &\leftrightarrow (V, E) \models g_n(y) \text{ is a subset of } x. \end{aligned}$$

Given $a \in V$, we thus infer that

$$\begin{aligned} (a, v) \in E &\leftrightarrow (a, g_{n+1}(A)) \in E \\ \leftrightarrow ((\exists \bar{a} \in A)(a = g_n(\bar{a})) &\leftrightarrow (\exists y \in U_n) a = g_n(y) \wedge f_{n+1}(y) \subset f_{n+1}(x) \\ \leftrightarrow (\exists z \in V) a = Z \wedge (V, E) \models z &\text{ is a subset of } v \\ \leftrightarrow (V, E) \models a &\text{ is a subset of } v. \end{aligned}$$

(5): Assume that $\varphi = \varphi(x, y) \in (\text{ZFC})$ and $u, y \in V$. Assume also that $u = g_{n+1}(x)$ and put $A := \{z \in f_{n+1}(x) : \varphi(g_n(z), y)\}$. Clearly, $A \in U_{n+1}$. Assign $v := g_{n+1}(A)$. In this event

$$\begin{aligned} (a, v) \in E &\leftrightarrow (\exists z \in U_n) a = g_n(z) \wedge z \in A \\ \leftrightarrow (\exists z \in U_n) z \in f_{n+1}(x) \wedge a = g_n(z) \wedge \varphi(g_n(z), y) \\ &\leftrightarrow (a, u) \in E \wedge \varphi(a, y) \end{aligned}$$

for $a \in U$.

(6): Assume that $u = g_{n+1}(x)$ and put

$$A := \{z \in U_n : z \in f_{n+1}(x) \wedge f_{n+1}(z) \neq \emptyset\}.$$

There is a choice function $\psi : A \rightarrow U_{n+1}$ such that $\psi(z) \in f_{n+1}(z)$ for all $z \in A$. The set ψ is a member of U_{n+3} . Put $f := g_{n+3}(\psi)$. It is easy to see that f plays the role of a function inside (V, E) , satisfying

$$(\forall v \in V)(v, y) \in E \wedge v \neq 0 \rightarrow (f(v), y) \in E.$$

It is such an element f that we have to demonstrate.

(7)–(10): These are beyond a doubt.

(11): Take $g_1(\overline{U})$ as \overline{U} (which is sound since $\overline{U} \in U_1$). By (4) $(v, \overline{U}) \in \iota \leftrightarrow (\exists u \in U_0)v = \iota(u) \wedge u \in \overline{U}$. In other words, $(v, \overline{U}) \in E \leftrightarrow v \in \iota(\overline{v})$. \triangleright

3.2.21. COMMENTS.

(1) Zermelo suggested in 1908 an axiomatics that coincides practically with ZF_1 – ZF_3 , ZF_5 , 3.2.5, and 3.2.6. This system, together with the Russell theory of types, is listed among the first formal axiomatics for set theory.

The axioms of extensionality ZF_1 and union ZF_2 were proposed earlier by Frege (1883) and Cantor (1899). The idea of the axiom of infinity ZF_5 belongs to Dedekind.

(2) The axiom of choice AC seems to be in use implicitly for a long time before it was distinguished by Peano in 1890 and Levy in 1902. This axiom was formally introduced by Zermelo in 1904 and remained most disputable for many years. The axiom of choice is part and parcel of the most vital fragments of contemporary mathematics. So, it is no wonder that AC is accepted by the overwhelming majority of working mathematicians. Discussions of the place and role of the axiom of choice may be found elsewhere [62, 123, 130, 196, 316].

(3) The axiomatics of ZFC was completely elaborated at the beginning of the 1920s. By that time the formalization of the set-theoretic language had been completed, which made it possible to clarify the vague description of the type of properties admissible in the axiom of comprehension. On the other hand, the Zermelo axioms do not yield the claim of Cantor that each bijective image of a set is a set. This drawback was obviated by Fraenkel in 1922 and Scolem in 1923 who suggested variations of the axiom of replacement. This moment seems to pinpoint the birth of ZFC.

(4) The axiom of regularity ZF_6 was in fact suggested by von Neumann in 1925. This axiom is independent of the other axioms of ZFC.

(5) The system of axioms of ZFC is infinite as noted in 3.2.4. Absence of finite axiomatizability for ZFC was proven by Montague in 1960 (see [123, 155, 316, 516]).

3.3. Nelson Internal Set Theory

The preliminary analysis of the properties of standard and nonstandard sets has shown that the von Neumann universe, if furnished with the predicate “to be or nor to be standard,” has ample room for infinitesimals but fails to accommodate their monad, them all “viewed as one.” In other words, infinitesimal analysis demonstrates that Zermelo–Fraenkel set theory, describing the classical world of “standard” mathematics, distinguishes only some internal part of the cantorion paradise, the universe of naive sets.

To emphasize this peculiarity, the nonstandard theory of sets we start presenting refers to the elements of the von Neumann universe as *internal sets*. Therefore, a set in the sense of Zermelo–Fraenkel set theory and an internal set are synonyms.

A convenient foundation for infinitesimal analysis is given by the so-called *internal set theory* suggested by Nelson and abbreviated to IST.

3.3.1. The alphabet of IST results from the alphabet of ZFC by adjoining the only one new symbol of the unary predicate St expressing the property of a set to be or not to be standard.

In other words, the texts of IST may contain fragments of the type $\text{St}(x)$ which reads as “ x is standard” or “ x is a standard set.” Therefore, the semantic domain of definition for variables of IST is the world of Zermelo–Fraenkel set theory; i.e., the von Neumann universe in which we can now discriminate between standard and nonstandard sets.

3.3.2. The formulas of IST are defined routinely on enriching the collection of atomic formulas with the texts like $\text{St}(x)$ where x is a variable. Each formula of ZFC is a formula of IST, whereas the converse is obviously false. To discriminate between the formulas of IST, we call the formulas of ZFC *internal*. By an *external* formula, we mean a formula IST not expressible in ZFC. For instance, the text “ x is standard” is an external formula of IST.

For ecological reasons, we use the following convenient abbreviations: To signify that φ is a formula of IST, we write $\varphi \in (\text{IST})$; similarly, $\varphi \in (\text{ZFC})$ means that φ is an internal formula of IST.

3.3.3. The distinction between the formulas of IST leads naturally to the notions of external and internal classes. If φ is an external formula of IST then we read the text $\varphi(y)$ as follows: “ y is an *element of the external class* $\{x : \varphi(x)\}$.” The term *internal class* implies the same as the term *class* in ZFC. We usually speak simply of classes when this leads to no confusion.

3.3.4. An external class consisting of elements of some internal set x is an *external set* or, in more detail, an external subset of x . It is worth noting that each internal class consisting of elements of some internal set is an internal set too.

Alongside the abbreviations of ZFC, the theory IST maintains some additional agreements. We list a few of them are concluded. A list of them follows:

$$\begin{aligned}
\mathbf{V}^{\text{st}} &:= \{x : \text{St}(x)\} \text{ is the external class of standard sets;} \\
x \in \mathbf{V}^{\text{st}} &:= x \text{ is standard} := (\exists y)(\text{St}(y) \wedge y = x); \\
(\forall^{\text{st}} x) \varphi &:= (\forall x)(x \text{ is standard} \rightarrow \varphi); \\
(\exists^{\text{st}} x) \varphi &:= (\exists x)(x \text{ is standard} \wedge \varphi); \\
(\forall^{\text{st fin}} x) \varphi &:= (\forall^{\text{st}} x)(x \text{ is finite} \rightarrow \varphi); \\
(\exists^{\text{st fin}} x) \varphi &:= (\exists^{\text{st}} x)(x \text{ is finite} \wedge \varphi); \\
{}^\circ x &:= \{y \in x : y \text{ is standard}\}.
\end{aligned}$$

The external set ${}^\circ x$ is often called the *standard part* or *standard core* of x . Observe conspicuous collision of notations and outright abuse of the language since ${}^\circ x$, with $x \in {}^\circ \mathbb{R}$, denotes the value of the standard part operation at x as well as the standard part of x as a set. Fortunately, these convenient collision and abuse are harmless, leading to confusion rarely if ever.

3.3.5. The list of axioms of IST includes all those of ZFC and three new axiom-schemata which are known collectively as the *principles of nonstandard set theory* or, which slightly abuses the language, the *principles of infinitesimal analysis*.

(1) Transfer Principle:

$$\begin{aligned}
&(\forall^{\text{st}} x_1)(\forall^{\text{st}} x_2) \dots (\forall^{\text{st}} x_n)((\forall^{\text{st}} x) \varphi(x, x_1, \dots, x_n)) \\
&\quad \rightarrow (\forall x) \varphi(x, x_1, \dots, x_n)
\end{aligned}$$

for every internal formula φ ;

(2) Idealization Principle:

$$\begin{aligned}
&(\forall x_1)(\forall x_2) \dots (\forall x_n)((\forall^{\text{st fin}} z)(\exists x)(\forall y \in z) \varphi(x, y, x_1, \dots, x_n)) \\
&\quad \leftrightarrow (\exists x)(\forall^{\text{st}} y) \varphi(x, y, x_1, \dots, x_n),
\end{aligned}$$

where $\varphi \in (\text{ZFC})$ is an arbitrary internal formula;

(3) Standardization Principle:

$$(\forall x_1) \dots (\forall x_n)((\forall^{\text{st}} x)(\exists^{\text{st}} y)(\forall^{\text{st}} z)(z \in y) \leftrightarrow (z \in x \wedge \varphi(z, x_1, \dots, x_n)))$$

for every (possibly external) formula φ .

3.3.6. Powell Theorem. IST is conservative over ZFC.

◁ Assume that φ is a formula of ZFC, $\varphi = \varphi(x_1, \dots, x_n)$, and φ is proven in IST. Assume further that the proof of φ uses the axioms ψ_1, \dots, ψ_m of ZFC. By the Montague–Levy Theorem there is an ordinal α such that

$$\varphi(x_1, \dots, x_n) \leftrightarrow \varphi^{V^\alpha}(x_1, \dots, x_n) \wedge \psi_1^{V^\alpha} \wedge \dots \wedge \psi_m^{V^\alpha}$$

for all $x_1, \dots, x_n \in V^\alpha$.

Put $U_0 := V^\alpha$ and $\varepsilon_0 := \in |_{V^\alpha \times V^\alpha}$. Take $\mathcal{P}_{\text{fin}}(U_0)$ as \mathcal{E} and let \mathcal{F} be some ultrafilter finer than the tail filter of $\mathcal{P}_{\text{fin}}(U_0)$. Denote the resulting ultrapower (= enlargement) of U_0 by U_1 and let $\iota_1 : U_0 \rightarrow U_1$ stand for the canonical embedding of U_0 into U_1 . Repeating this construction successively, put $U_{n+1} := U_n^{\mathcal{P}_{\text{fin}}(U_0) / \mathcal{F}}$ and let $\iota_{n+1} : U_n \rightarrow U_{n+1}$ be the canonical embedding of U_n into the ultrapower U_{n+1} . Identifying U_n and $\iota_{n+1}(U_n)$, consider U_n as a subset of U_{n+1} . Put $U := \bigcup_{n \in \mathbb{Z}} U_n$ and $\varepsilon := \bigcup_{n \in \mathbb{Z}} \varepsilon_n$, where $\varepsilon_{n+1} := \iota_n \circ \varepsilon_n \circ \iota_{n+1}$ is the appropriate interpretation of membership. Assume further that $* : U_0 \rightarrow U$ stands for the canonical embedding of U_0 into U . Treat the predicate $\text{St}(\cdot)$ as membership in $\{*u : u \in U_0\}$. Since $x \in V^\alpha \rightarrow (\exists \beta \in \alpha) x \in V^\beta \rightarrow \mathcal{P}(x) \in V^\beta$, conclude that the standardization principle holds in U . Validity of the principles of transfer and idealization follows from the Łoś Theorem. Therefore, $\psi_1^s, \dots, \psi_m^s$ and the principles of IST are satisfied in (U, ε) , implying that $\varphi^U(*x_1, \dots, *x_n)$ and $\varphi^{V^\alpha}(x_1, \dots, x_n)$. Consequently, φ holds in the von Neumann universe. ▷

3.3.7. The above theorem implies that the internal theorems of IST are theorems of ZFC. In other words, proving “standard” theorems on sets we may use the formalism of IST with the same feeling of reliability as we enjoy in ZFC. At the same time, we should always bear in mind that the ultimate foundation ZFC resides in the long run in its practical infallibility and semantic justification.

3.3.8. Pondering over the meaning of the formal records of the axioms of IST, we cannot help but notice that the idealization principle looks somewhat cumbersome. While the principles of transfer and standardization in their formal disguise adequately reflect the naive conception, put forward earlier, the formal expression of the idealization principle may drive the reader in a quandary. To clear the coast, we start with proving that idealization principle as stated in 3.3.5 (2) guarantees the presence of nonstandard elements.

3.3.9. *There is a finite internal set containing every standard set.*

◁ Consider the following formula: $\varphi := (x \text{ is finite} \wedge (y \in x))$. Note that $\varphi \in (\text{ZFC})$. Given a standard finite z , we may find an element x such that $\varphi(x, y)$ for all $y \in z$. Indeed, z may serve as such an x . We then complete the proof by idealization. ▷

3.3.10. When applying the idealization principle, we should bear in mind clearly that those finite sets are standard whose every element is standard. This fact was proved earlier, cf. 2.2.2. It is instructive to inspect the available formal inference which proceeds also by idealization.

3.3.11. *If A is an internal set then*

$$A = {}^\circ A \leftrightarrow (A \text{ is standard}) \wedge (A \text{ is finite}).$$

◁ Look at the formula $\varphi := x \in A \wedge x \neq y$. Clearly, $\varphi \in (\text{ZFC})$. Proceed by idealization as follows:

$$\begin{aligned} (\forall^{\text{st fin}} z)(\exists x)(\forall y \in z) \varphi(x, y, A) &\leftrightarrow (\exists x)(\forall^{\text{st}} y)(x \in A \wedge x \neq y) \\ &\leftrightarrow (\exists x \in A)(x \text{ is nonstandard}) \leftrightarrow A - {}^\circ A \neq \emptyset. \end{aligned}$$

In other words,

$$\begin{aligned} A = {}^\circ A &\leftrightarrow (\exists^{\text{st fin}} z)(\forall x)(\exists y \in z) x \notin A \vee x = y \\ &\leftrightarrow (\exists^{\text{st fin}} z)(\forall x \in A)(\exists y \in z) x = y \leftrightarrow (\exists^{\text{st fin}} z) A \subset z, \end{aligned}$$

which completes the proof. ▷

3.3.12. Construction Principle. *Assume that X and Y are standard sets, while $\varphi = \varphi(x, y, z)$ is a formula of IST. Then*

$$\begin{aligned} &(\forall^{\text{st}} x)(\exists^{\text{st}} y)(x \in X \rightarrow y \in Y \wedge \varphi(x, y, z)) \\ &\leftrightarrow (\exists^{\text{st}} y(\cdot))(\forall^{\text{st}} x)(y(\cdot) \text{ is a function from } X \text{ in } Y \\ &\quad \wedge (x \in X \rightarrow \varphi(x, y(x), z))). \end{aligned}$$

◁ Consider the standardization $\overline{F}(x) := * \{y \in Y : \varphi(x, y, z)\}$. Applying 3.3.5 (3) again, arrange the standard set

$$F := * \{(x, A) \in X \times \mathcal{P}(Y) : \overline{F}(x) = A\},$$

which is sound since $\mathcal{P}(Y)$ is standard whenever so is Y . By hypothesis, $(\forall^{\text{st}} x \in X) \overline{F} \neq \emptyset$ and so. In this case, $F(x) = \overline{F}(x)$ by the definition of F . Consequently,

$$((\forall^{\text{st}} x \in X)(F(x) \neq \emptyset)) \rightarrow ((\forall x \in X)(F(x) \neq \emptyset))$$

by transfer. Appealing now to the axiom of choice, conclude that

$$(\exists y(\cdot))(y(\cdot) \text{ is a function from } X \text{ to } Y) \wedge (\forall x \in X)(y(x) \in F(x)).$$

By transfer, there is a standard function $y(\cdot)$ from X to Y satisfying $y(x) \in F(x)$ for all $x \in X$. Recalling the definition of F once again, observe that $y(\cdot)$ is a sough function. ▷

3.3.13. In what follows (as well as precedes) we find it convenient to use some symbolic records of the rules collected in the above propositions, deliberately but slightly abusing the language. For instance, the construction principle 3.3.12 may be rewritten as

- (1) $(\forall^{\text{st}}x)(\exists^{\text{st}}y) \varphi(x, y) \leftrightarrow (\exists^{\text{st}}y(\cdot))(\forall^{\text{st}}x) \varphi(x, y(x)),$
 (2) $(\exists^{\text{st}}x)(\forall^{\text{st}}y) \varphi(x, y) \leftrightarrow (\forall^{\text{st}}y(\cdot))(\exists^{\text{st}}x) \varphi(x, y(x)),$

where $\varphi \in (\text{IST})$. In other words, we will neglect the possible presence of unbound variables in φ and the assumption that all quantifiers must be bounded which implies that x and y are assumed to range over some standard sets that are specified in advance. Similarly, if $\varphi = \varphi(x_1, \dots, x_n)$ and $\psi = \psi(y_1, \dots, y_n)$ then we will write $\varphi \leftrightarrow \psi$ whenever

$$(\forall^{\text{st}}x_1) \dots (\forall^{\text{st}}x_n)(\forall^{\text{st}}y_1) \dots (\forall^{\text{st}}y_n) \varphi(x_1, \dots, x_n) \leftrightarrow \psi(y_1, \dots, y_n),$$

and say that φ and ψ are equivalent (although if one of the formulas φ and ψ is external then the formulas $\varphi(x_1, \dots, x_n)$ and $\psi(y_1, \dots, y_n)$ are not equivalent for some assignment of variables). Using these agreements, we will express the transfer principle in reduced form

- (3) $(\forall^{\text{st}}x) \varphi(x) \leftrightarrow (\forall x) \varphi(x),$
 (4) $(\exists^{\text{st}}x) \varphi(x) \leftrightarrow (\exists x) \varphi(x),$

always keeping in mind that φ must be an internal formula: $\varphi \in (\text{ZFC})$.

It is reasonable also to write down a few elementary rules valid for every φ :

- (5) $(\forall x)(\forall^{\text{st}}y) \varphi(x, y) \leftrightarrow (\forall^{\text{st}}y)(\forall x) \varphi(x, y),$
 (6) $(\exists x)(\exists^{\text{st}}y) \varphi(x, y) \leftrightarrow (\exists^{\text{st}}y)(\exists x) \varphi(x, y),$

as well as some new records of the idealization principle:

- (7) $(\forall^{\text{st fin}}z)(\exists x)(\forall y \in z) \varphi(x, y) \leftrightarrow (\exists x)(\forall^{\text{st}}y) \varphi(x, y),$
 (8) $(\exists^{\text{st fin}}z)(\forall x)(\exists y \in z) \varphi(x, y) \leftrightarrow (\forall x)(\exists^{\text{st}}y) \varphi(x, y),$

which apply, obviously, only $\varphi \in (\text{ZFC})$.

3.3.14. These rules allow us to translate many (but definitely not all) concepts and statements of infinitesimal analysis into equivalent definitions and theorems free from the notion of “standardness.” In other words, the formulas of IST expressing “something unusual” about standard objects can be translated into equivalent formulas of ZFC which are in mathematical parlance. The procedure yielding this result is the *Nelson algorithm* or *reduction algorithm* which rests essentially on 3.3.13 (1)–(8). The essence of this “decoding” algorithm is in using standard functions, idealization and transposition of quantifiers for reducing the expression under study to a form more suitable for transfer. This translation amounts ultimately to *elimination of the external notion of standardness*. It is worth observing that each case of application of the formulas of 3.3.13 requires ensuring the hypotheses of their validity.

3.3.15. The Nelson algorithm consists of the following steps:

- (1) Rewrite a proposition of infinitesimal analysis as a formula ψ of IST; i.e., decode all abbreviations;
- (2) Reduce the formula ψ of step (1) to the prenex normal form

$$(Q_1x_1) \dots (Q_nx_n) \varphi(x_1, \dots, x_n),$$

where $Q_k \in \{\forall, \exists, \forall^{\text{st}}, \exists^{\text{st}}\}$ for $k := 1, \dots, n$ and the matrix φ of ψ is a formula of ZFC;

- (3) In case Q_n is an “*internal*” quantifier, i.e., \forall or \exists ; proceed by putting $\varphi := (Q_nx_n) \varphi(x_1, \dots, x_n)$ and return to step (2);
- (4) If Q_n is an “*external*” quantifier, i.e., \forall^{st} or \exists^{st} ; search the prefix $(Q_1x_1) \dots (Q_nx_n)$ from right to left until the first internal quantifier is found;
- (5) In case there is no internal quantifiers in step (4); replace the quantifier Q_n with the corresponding internal quantifier (by 3.3.13 (3) and 3.3.13 (4)), and return to step (2), i.e., delete all superscripts st in right-to-left order;
- (6) Let Q_m be the first internal quantifier encountered. Suppose that Q_{m+1} is an external quantifier of the same type as Q_m (i.e., $Q_m = \forall$ and $Q_{m+1} = \forall^{\text{st}}$, or $Q_m = \exists$ and $Q_{m+1} = \exists^{\text{st}}$). Transpose the quantifiers by 3.3.13 (5) and 3.3.13 (6) and return to step (2);
- (7) In case all the quantifiers Q_{m+1}, \dots, Q_n are of the same type, apply the idealization principle in the form 3.3.13 (7) or 3.3.13 (8), and return to step (2);
- (8) In case the quantifiers alternate; i.e., Q_{p+1} is of the same type as Q_m while all quantifiers Q_{m+1}, \dots, Q_p are of the opposite type, apply 3.3.13 (1) or 3.3.13 (2), on assuming $x := (x_{m+1}, \dots, x_p)$ and $y := x_{p+1}$. After that return to step (2).

3.3.16. Note that an arbitrary assertion is expressible in various forms some of which can be absolutely incomprehensible. Hence, implementing the Nelson algorithm in *practice*, we must seize all possibilities of accelerating the procedure of “dragging out the external quantifiers.” In particular, it is wise sometimes to skip or leave unfinished step 3.3.15 (2). We will demonstrate this by example.

3.3.17. EXAMPLES.

- (1) *Infinitesimal analysis enjoys the external induction principle:*

$$(\varphi(1) \wedge ((\forall n \in {}^\circ\mathbb{N})(\varphi(n) \rightarrow \varphi(n+1)))) \rightarrow (\forall n \in {}^\circ\mathbb{N}) \varphi(n)$$

for every $\varphi \in (\text{IST})$.

◁ The formula φ may be external, which prevents us from applying the Nelson algorithm directly to the formal record of the external induction principle.

So, consider the standardization $A := *\{n \in \mathbb{N} : \varphi(n)\}$. Obviously, $1 \in A$, and $n + 1 \in A$ for every standard $n \in A$.

We must prove that ${}^\circ\mathbb{N} \subset A$. For this purpose, write down the formula in question and apply the Nelson algorithm to it:

$$\begin{aligned} & (1 \in A \wedge (\forall^{\text{st}} n \in \mathbb{N})(n \in A \rightarrow (n + 1) \in A)) \rightarrow {}^\circ\mathbb{N} \subset A \\ \leftrightarrow & (\forall^{\text{st}} m)(\forall^{\text{st}} n)(m \in \mathbb{N} \wedge n \in \mathbb{N} \wedge 1 \in A \wedge n \in A \rightarrow (n + 1) \in A) \\ \rightarrow & m \in A \leftrightarrow (1 \in A \wedge (\forall n \in \mathbb{N})(n \in A \rightarrow (n + 1) \in A)) \rightarrow \mathbb{N} \subset A, \end{aligned}$$

which completes the proof. ▷

(2) *The sum of infinitesimals is an infinitesimal too.*

$$\begin{aligned} & \triangleleft (\forall s \in \mathbb{R})(\forall t \in \mathbb{R})(s \approx 0 \wedge t \approx 0 \rightarrow s + t \approx 0) \\ \leftrightarrow & (\forall s \in \mathbb{R})(\forall t \in \mathbb{R})(s \approx 0 \wedge t \approx 0 \rightarrow (\forall^{\text{st}} \varepsilon > 0)(|s + t| < \varepsilon)) \\ \leftrightarrow & (\forall^{\text{st}} \varepsilon > 0)(\forall s \in \mathbb{R})(\forall t \in \mathbb{R})(\forall^{\text{st}} \delta_1 > 0) \\ & (\forall^{\text{st}} \delta_2 > 0)(|s| < \delta_1 \wedge |t| < \delta_2 \rightarrow |s + t| < \varepsilon) \\ \leftrightarrow & (\forall^{\text{st}} \varepsilon > 0)(\forall s \in \mathbb{R})(\forall t \in \mathbb{R})(\exists^{\text{st}} \delta_1 > 0)(\exists^{\text{st}} \delta_2 > 0)(|s| < \delta_1 \\ & \wedge |t| < \delta_2 \rightarrow |s + t| < \varepsilon) \leftrightarrow (\forall^{\text{st}} \varepsilon)(\forall s)(\forall t)(\exists^{\text{st}} \delta_1)(\exists^{\text{st}} \delta_2)(\varepsilon > 0 \\ & \wedge \dots \wedge \delta_2 > 0 \wedge |s| < \delta_1 \wedge |t| < \delta_2 \rightarrow |s + t| < \varepsilon) \\ \leftrightarrow & (\forall^{\text{st}} \varepsilon)(\forall s)(\forall t)(\exists^{\text{st}} \delta_1)(\exists^{\text{st}} \delta_2)(|s| < \delta_1 \wedge |t| < \delta_2 \rightarrow |s + t| < \varepsilon) \\ \leftrightarrow & (\forall^{\text{st}} \varepsilon)(\exists^{\text{st fin}} \Delta_1)(\exists^{\text{st fin}} \Delta_2)(\forall s)(\forall t)(\exists \delta_1 \in \Delta_1)(\exists \delta_2 \in \Delta_2) \\ & (|s| < \delta_1 \wedge |t| < \delta_2 \rightarrow |s + t| < \varepsilon) \\ \leftrightarrow & (\forall^{\text{st}} \varepsilon)(\exists^{\text{st}} \delta_1)(\exists^{\text{st}} \delta_2)(\forall |s| < \delta_1)(\forall |t| < \delta_2)(|s + t| \leq \varepsilon) \\ \leftrightarrow & (\forall \varepsilon > 0)(\exists \delta > 0)(\forall |s| < \delta)(\forall |t| < \delta)(|s + t| \leq \varepsilon). \quad \triangleright \end{aligned}$$

(3) Robinson Lemma. *Let (a_n) be an internal numerical sequence, and $a_n \approx 0$ for all $n \in {}^\circ\mathbb{N}$. Then there is an index $N \approx +\infty$ such that $a_n \approx 0$ for all $n \leq N$.*

◁ Apply the Nelson algorithm to the conclusion:

$$\begin{aligned}
& (\exists N \approx +\infty)(\forall n \leq N)(a_n \approx 0) \\
& \leftrightarrow (\exists N \in \mathbb{N})(\forall^{\text{st}} m \in \mathbb{N})((N \geq m) \wedge (\forall n \in \mathbb{N})(n \leq N)) \\
& \quad \rightarrow (\forall^{\text{st}} \varepsilon > 0)|a_n| < \varepsilon \leftrightarrow (\exists N)(\forall^{\text{st}} m)(\forall^{\text{st}} \varepsilon) \\
& \quad (\forall n)(N \geq m \wedge (n \leq N \rightarrow |a_n| < \varepsilon)) \\
& \leftrightarrow (\forall^{\text{st}} \{m_1, \dots, m_p\})(\forall^{\text{st}} \{\varepsilon_1, \dots, \varepsilon_p\})(\exists N)(\forall k := 1, \dots, p) \\
& \quad (N \geq m_k \wedge n \leq N \rightarrow |a_n| < \varepsilon_k) \\
& \leftrightarrow (\forall^{\text{st}} m)(\forall^{\text{st}} \varepsilon)(\exists N)(N \geq m \wedge (n \leq N \rightarrow |a_n| < \varepsilon)) \\
& \leftrightarrow (\forall^{\text{st}} m)(\forall^{\text{st}} \varepsilon)(m \in \mathbb{N} \wedge \varepsilon > 0 \rightarrow |a_m| < \varepsilon).
\end{aligned}$$

Proceed with applying the Nelson algorithm to the premise of the claim:

$$\begin{aligned}
& (\forall n \in {}^\circ\mathbb{N})(a_n \approx 0) \leftrightarrow (\forall^{\text{st}} n)(n \in \mathbb{N} \rightarrow (\forall^{\text{st}} \varepsilon > 0)(|a_n| < \varepsilon)) \\
& \leftrightarrow (\forall^{\text{st}} n)(\forall^{\text{st}} \varepsilon)(n \in \mathbb{N} \wedge \varepsilon > 0 \rightarrow |a_n| < \varepsilon).
\end{aligned}$$

Clearly, the premise and conclusion are equivalent. ▷

(4) Unique Determination Principle. *Unique determination implies standardness in the standard environment.*

In symbols, if $y, V \in \mathbf{V}^{\text{st}}$ and $\varphi = \varphi(x, y)$ is an arbitrary (possibly external) formula of IST then

$$((\exists! \bar{x} \in V) \varphi(\bar{x}, y)) \rightarrow \text{St}(\bar{x}).$$

◁ Using the Nelson algorithm, rewrite φ as

$$\varphi(x, y) := (\forall^{\text{st}} u)(\exists^{\text{st}} v) \psi(x, u, v, y),$$

with $\psi \in (\text{ZFC})$.

In particular, by the construction principle,

$$(\exists^{\text{st}} \bar{v}(\cdot))(\forall^{\text{st}} u) \psi(\bar{x}, u, \bar{v}(u), y). \quad (1)$$

Moreover,

$$(\forall z)(\forall^{\text{st}} u \exists^{\text{st}} v \psi(z, u, v, y) \rightarrow z = \bar{x}).$$

Using the Nelson algorithm, infer

$$(\forall^{\text{st}} v(\cdot))(\exists^{\text{st fin}} U)(\forall z)((\forall u \in U) \psi(z, u, v(u), y) \rightarrow z = \bar{x}). \quad (2)$$

Denote by \bar{U} the standard finite set that corresponds by (2) to the function $\bar{v}(\cdot)$ of (1).

Clearly, $(\forall u \in \bar{U}) \psi(x, u, \bar{v}(u), y)$ and so $(\exists z)(\forall u \in \bar{U}) \psi(z, u, \bar{v}(u), y)$. By transfer,

$$(\exists^{\text{st}} z)(\forall u \in \bar{U}) \psi(z, u, v(u), y).$$

Using (2), conclude that $z = \bar{x}$, i.e., $\text{St}(\bar{x})$. ▷

3.3.18. Alongside the theory IST by Nelson we will use some variation of this theory, the so-called bounded (or limited) set theory BST, that was suggested by Kanoveř and Reeken in [213].

The theory BST preserves the alphabet and related attributes of IST. Distinctions appear only in the formulation of the principles of nonstandard set theory. Moreover, the principles of transfer and standardization of IST are still valid in BST. However, the theory BST considers each internal set as included in some standard set, implying due restrictions on the process of idealization:

(1) Boundedness Principle:

$$(\forall x)(\exists^{\text{st}} X)(x \in X);$$

(2) Bounded Idealization Principle:

$$\begin{aligned} & (\forall x_1)(\forall x_2) \dots (\forall x_n)((\forall^{\text{st fin}} z)(\exists x \in z_0)(\forall y \in z) \\ & \varphi(x, y, x_1, \dots, x_n) \leftrightarrow (\exists x \in z_0)(\forall^{\text{st}} y) \varphi(x, y, x_1, \dots, x_n)), \end{aligned}$$

where $\varphi \in (\text{ZFC})$ is an internal formula and z_0 is some standard set.

Kanoveř and Reeken proved that the bounded idealization principle may be replaced with the following

(3) Internal Saturation Principle:

$$\begin{aligned} & (\forall x_1)(\forall x_2) \dots (\forall x_n)((\forall^{\text{st fin}} z \subset z_0) (\exists x)(\forall y \in z) \varphi(x, y, x_1, \dots, x_n) \\ & \leftrightarrow (\exists x)(\forall^{\text{st}} y \in z_0) \varphi(x, y, x_1, \dots, x_n)), \end{aligned}$$

where $\varphi \in (\text{ZFC})$ is an internal formula and z_0 is some standard set.

The most important property of BST is the fact that the *bounded* or *limited* universe, comprising the members of all standard set in IST, serves as a model for BST. This implies that BST is conservative over ZFC.

3.4. External Set Theories

The basic principles of infinitesimal analysis are adequately reflected in the formal apparatus of Nelson internal set theory. The Powell Theorem enables us to view IST as a technique of studying the von Neumann universe. At the same time, the presence of external objects completely undermines the popular belief that the formalism of Zermelo–Fraenkel set theory brings about a sufficient freedom of operation from the viewpoint of the naive set theory.

Residing in the realm of IST, we are not in a position even to raise such an innocent question as: “Is it possible to distinguish some set of reals such that each element of \mathbb{R} will admit a unique expansion in a linear combination of them with

STANDARD coefficients, since \mathbb{R} may be naturally viewed as a vector space over ${}^\circ\mathbb{R}$?" The amount of these inadmissible questions, undoubtedly mathematical from the semantic point of view, is great to the extent that makes the necessity of trespassing the frontiers of IST vital.

A priori prohibition against formulating problems is nothing but blindfolding the mind. The introduction of an ad hoc dogma, an "explicitly expressed prohibition against thinking" (as was aphoristically remarked by Feuerbach) is absolutely unacceptable in searching truth. A practical solution of the problem of returning to the cantorion paradise lies, in particular, in designing a formalism that would allow us to use the conventional means for working with sets that are external to the von Neumann universe. We now start reviewing some axiomatic approaches to external sets.

The first version of an appropriate formalism was suggested by Hrbáček in his *external set theory* abbreviated as EXT. A close version, *nonstandard set theory* NST, was propounded later by Kawai. Semantically speaking, these set theories demonstrate that the external universe is constructed in much the same way as the universe of naive sets as viewed from the standpoint of a pragmatic philistine which the working mathematician often occupies and mostly indulges in. Clearly, *the external universe is at least as good as the naive cantorion paradise* since the former admits all classical set-theoretic operations including the generation of subsets by abstract predicates (the axioms of comprehension) as well as the possibility of well-ordering arbitrary sets (the axiom of choice). At the same time the external universe includes the full collection of all internal sets, standard and nonstandard, satisfying the principles of transfer, idealization, and standardization in the forms close to their intuitive content. Strictly speaking, all internal sets belong to the external world by definition.

As regards the every-day needs of the conventional (standard and nonstandard) mathematical analysis, the two theories EXT and NST supply the tools of practically the same power in the quantities more than enough for meeting the demands of whatever version of analysis.

Of course, to peruse the details of the formal axiomatics of both theories is necessary for avoiding the illusions that accompany the euphoria of universal permissiveness. It is worth noting that *the external world does not coincide with the von Neumann universe* (the axiom of regularity is omitted, which is essential). Moreover, the exact formulations of the principles of infinitesimal analysis of EXT differ in technical details from their analogs in IST. Therefore, EXT is not an extension of IST whereas EXT is a conservative extension of ZFC. This gap was filled by Kawai whose theory NST enriches the formal apparatus of IST and, alongside with IST and EXT, provides a reliable technique for studying ZFC.

3.4.1. The alphabet of EXT comprises the alphabet IST and a sole new symbol

of the unary predicate, Int , expressing the property of a set to be or not to be internal. In other words, we accept for consideration the texts that contain $\text{Int}(x)$, or in a more verbose form “ x is internal,” or amply “ x is an internal set.” Intuitively, the range of variables of EXT is the *external universe* $\mathbf{V}^{\text{Ext}} := \{x : x = x\}$ which includes the *internal universe* $\mathbf{V}^{\text{Int}} := \{x \in \mathbf{V}^{\text{Ext}} : \text{Int}(x)\}$ whose part is the *standard universe* $\mathbf{V}^{\text{St}} := \{x \in \mathbf{V}^{\text{Ext}} : \text{St}(x)\}$.

3.4.2. The conventions of EXT are analogous to those of ZFC and IST. In particular, we will continue using the “classifiers,” i.e., braces (see 3.3.3) and the traditional symbols for the simplest operations on classes of external sets. Given a formula φ of EXT (in symbols, $\varphi \in (\text{EXT})$) we proceed as before and put

$$\begin{aligned} (\forall^{\text{St}} x) \varphi &:= (\forall x)(\text{St}(x) \rightarrow \varphi) := (\forall x \in \mathbf{V}^{\text{St}}) \varphi, \\ (\exists^{\text{Int}} x) \varphi &:= (\exists x)(\text{Int}(x) \wedge \varphi) := (\exists x \in \mathbf{V}^{\text{Int}}) \varphi. \end{aligned}$$

Similar rules, easily understood from context, will be hereupon used without further specification. We also need a new concept and a corresponding notation. We will say that an external set A has *standard size* (in writing, $A \in V^{\text{size}}$) provided that there are a standard set a and an external function f satisfying $(\forall X)(X \in A \leftrightarrow (\exists^{\text{st}} x \in a) X = f(x))$.

3.4.3. Let $\varphi \in (\text{ZFC})$ be a formula of EXT that is also a formula of ZFC (i.e., φ contains no symbols St and Int). Replacing each occurrence of a quantifier Q , where $Q \in \{\forall, \exists\}$, in the record of φ by Q^{st} , denote the result by φ^{St} and call φ^{St} the *standardization* of φ or *relativization* of φ to \mathbf{V}^{St} . By analogy, replacing each occurrence of a quantifier Q with Q^{Int} , we come to the formula φ^{Int} which is called the *internalization* of φ or *relativization* of φ to \mathbf{V}^{Int} . Note that nothing happens to the unbound variables of φ .

We tacitly proceed likewise with abbreviations. For instance, given external sets A and B , we write

$$\begin{aligned} A \subset^{\text{St}} B &:= (\forall^{\text{st}} x)(x \in A \rightarrow x \in B) \\ &:= ((\forall x)(x \in A \rightarrow x \in B))^{\text{St}} := (A \subset B)^{\text{St}}; \\ A \in^{\text{Int}} B &:= (A \in B)^{\text{Int}} := A \in B := A \in^{\text{St}} B := (A \in B)^{\text{St}}. \end{aligned}$$

3.4.4. The special axioms of EXT fall into three groups: (a) the *axioms of external set formation*, (b) the *axioms of interplay* between the universes \mathbf{V}^{St} , \mathbf{V}^{Int} , and \mathbf{V}^{Ext} , and, finally, (3) the *principles of transfer, idealization, and standardization*.

3.4.5. The universe EXT obeys the laws of *Zermelo set theory* (in symbols, Z). The following axioms for external set formation accepted:

(1) **Axiom of Extensionality:**

$$(\forall A), (\forall B)((A \subset B \wedge B \subset A) \leftrightarrow A = B);$$

(2) **Axiom of Pairing:**

$$(\forall A)(\forall B) \{A, B\} \in \mathbf{V}^{\text{Ext}};$$

(3) **Axiom of Union:**

$$(\forall A) \bigcup A \in \mathbf{V}^{\text{Ext}};$$

(4) **Axiom of Powersets:**

$$(\forall A) \mathcal{P}(A) \in \mathbf{V}^{\text{Ext}};$$

(5) **Axiom-Schema of Comprehension:**

$$(\forall A)(\forall X_1) \dots (\forall X_n) \{X \in A : \varphi(X, X_1, \dots, X_n)\} \in \mathbf{V}^{\text{Ext}}$$

for an arbitrary formula $\varphi \in (\text{EXT})$;

(6) **Axiom of Well-Ordering:** *Every external set is well-orderable.*

The last property, known also as the *Zermelo Theorem*, ensures (cf. (3.2.10)) the axiom of choice either in the conventional multiplicative form or in the form of the Kuratowski–Zorn Lemma. It is worth observing here that the routine list of the axioms of Z usually includes the axiom of infinity, which will appear somewhat later in EXT.

3.4.6. The second group of the axioms of EXT comprises the following:

(1) **Modeling Principle.** *The internal universe \mathbf{V}^{Int} is the von Neumann universe; i.e., if φ is an axiom of Zermelo–Fraenkel set theory then the internalization φ^{Int} is an axiom of EXT;*

(2) **Axiom of Transitivity:**

$$(\forall x \in \mathbf{V}^{\text{Int}}) x \subset \mathbf{V}^{\text{Int}},$$

i.e., internal sets are composed of only internal elements;

(3) **Axiom of Embedding:**

$$\mathbf{V}^{\text{St}} \subset \mathbf{V}^{\text{Int}},$$

i.e., standard sets are internal.

3.4.7. The third group of the axioms of EXT includes the following:

(1) Transfer Principle:

$$(\forall^{\text{st}} x_1) \dots (\forall^{\text{st}} x_n) \varphi^{\text{St}}(x_1, \dots, x_n) \leftrightarrow \varphi^{\text{Int}}(x_1, \dots, x_n)$$

for every formula $\varphi \in (\text{ZFC})$;

(2) Idealization Principle:

$$\begin{aligned} & (\forall^{\text{Int}} x_1) \dots (\forall^{\text{Int}} x_n) (\forall A \in V^{\text{size}}) ((\forall^{\text{fin}} z) z \subset A \\ & \quad \rightarrow (\exists^{\text{Int}} x) (\forall y \in z) \varphi^{\text{Int}}(x, y, x_1, \dots, x_n)) \\ & \rightarrow (\exists^{\text{Int}} x) (\forall^{\text{Int}} y \in A) \varphi^{\text{Int}}(x, y, x_1, \dots, x_n) \end{aligned}$$

for an arbitrary $\varphi \in (\text{ZFC})$;

(3) Standardization Principle:

$$(\forall A) (\exists^{\text{st}} a) (\forall^{\text{st}} x) (x \in A \leftrightarrow x \in a),$$

to each external set A there corresponds the standardization *A of A .

3.4.8. As a simplest useful corollary to the above axioms we mention that the *bounded formulas of ZFC are absolute*. To be more precise, if $\varphi \in (\Sigma_0)$ then

$$\begin{aligned} & (\forall^{\text{Int}} x_1) \dots (\forall^{\text{Int}} x_n) \varphi(x_1, \dots, x_n) \leftrightarrow \varphi^{\text{Int}}(x_1, \dots, x_n), \\ & (\forall^{\text{st}} x_1) \dots (\forall^{\text{st}} x_n) \varphi^{\text{St}}(x_1, \dots, x_n) \\ & \leftrightarrow \varphi^{\text{Int}}(x_1, \dots, x_n) \leftrightarrow \varphi(x_1, \dots, x_n). \end{aligned}$$

In other words, we may safely express every “bounded” property of standard sets both in terms of external and internal or standard elements. For instance, $x \subset y \leftrightarrow x \subset^{\text{St}} y \leftrightarrow x \subset^{\text{Int}} y$ for standard sets x and y .

3.4.9. Hrbáček Theorem. EXT is conservative over ZFC. In symbols,

$$\begin{aligned} & (\varphi \text{ is a theorem of ZFC}) \leftrightarrow (\varphi^{\text{Int}} \text{ is a theorem of EXT}) \\ & \leftrightarrow (\varphi^{\text{St}} \text{ is a theorem of EXT}) \end{aligned}$$

for all $\varphi \in (\text{ZFC})$.

◁ The bulky proof of this theorem can be found in [185]. ▷

3.4.10. Pondering over the above axiomatics, we must realize first of all that EXT is not an extension of IST.

In other words, the internal universe \mathbf{V}^{Int} is not a model for Nelson's IST since the idealization and standardization principles are formulated differently in EXT and IST.

The conditions for standardization in \mathbf{V}^{Int} are considerably less restrictive than those in IST. Indeed, given $\varphi \in (\text{IST})$ and $A \in \mathbf{V}^{\text{Int}}$, we may arrange $^*\{x \in A : \varphi(x)\}$ since $\{x \in A : \varphi(x)\}$ is an external subset of A . In IST this is possible provided that A is standard (do not forget that to standardize the collection of all standard elements is impossible in IST). Residing in EXT, we can find neither external nor internal set containing all standard sets. Indeed:

3.4.11. *There is no external set, a member of \mathbf{V}^{Ext} , containing every standard set.*

◁ Assume by contradiction that $\mathbf{V}^{\text{St}} \subset X$ for some $X \in \mathbf{V}^{\text{Ext}}$. Applying the axiom of comprehension 3.4.5 (5), to the formula $\varphi(x) = \text{St}(x)$ conclude that \mathbf{V}^{St} is an external set, i.e., $(\exists Y)(\forall Z)(Z \in Y \leftrightarrow \text{St}(Z))$. Look at the standardization $^*\mathbf{V}^{\text{St}}$. Observe that $^*\mathbf{V}^{\text{St}}$ is a standard finite set containing every standard set, which is impossible. ▷

3.4.12. Proposition 3.4.11 shows that the idealization principle in EXT (if relativized to \mathbf{V}^{Int}) differs from its counterpart in IST not only in form but also in essence. However, the importance of these differences should not be exaggerated.

3.4.13. *The following hold:*

- (1) *External naturals are the same as standard naturals;*
- (2) *A finite external set A is standard if and only if A comprises only standard elements;*
- (3) *The standard part ${}^\circ A := \{a \in A : \text{St}(a)\}$ of each external set A is of standard size;*
- (4) *Each infinite internal set contains a nonstandard element.*

◁ (1): By induction on standard naturals (which is, obviously, legitimate in EXT (cf. 2.2.2 (1))), $\mathbb{N}^{\text{Ext}} \supset {}^\circ\mathbb{N}$ with \mathbb{N}^{Ext} standing for the external set of external naturals. It is also clear that $^*\emptyset = \emptyset$ and $^*1 = ^*\{\emptyset\} = \{\emptyset\} = 1$. Hence, by induction on external naturals (which is a routine theorem of Z), $\mathbb{N}^{\text{Ext}} \subset {}^\circ\mathbb{N}$ and so ${}^\circ\mathbb{N} = \mathbb{N}^{\text{Ext}}$.

(2): Every standard set is internal. We may thus proceed along the lines of the proof of 2.2.2 (3) on appealing to 3.4.6 (2). Also, by 2.2.2 (2), a finite set composed of standard elements is standard.

(3): Let *A stand for the standardization of A . Put $f(a) := a$ for $a \in {}^\circ A$. Obviously, $(\forall X)(X \in {}^\circ A \leftrightarrow (\exists^{\text{st}} x \in ^*A) f(x) = X)$.

(4): Let A stand for the internal set under discussion. By (3), the set ${}^{\circ}A$ is of standard size. Hence, we may apply the idealization principle for $\varphi(x, y) := y \neq x \wedge x \in A$. Since A is infinite; therefore, $(\exists x \in A)(\forall y \in z) x \neq y$ for every finite $z \subset {}^{\circ}A$. Finally, $(\exists x \in A)(\forall y \in {}^{\circ}A) x \neq y$. \triangleright

3.4.14. As prompted by 3.4.13 and 3.4.9, it is convenient to consider some variation INT of IST which is a conservative extension of ZFC with EXT serving in turn as an extension of INT. The difference between INT and IST lies in the idealization and standardization principles which are stated in INT as follows:

- (1) $(\forall A)(\forall x_1) \dots (\forall x_n)((\forall^{\text{st fin}} z)(z \subset A)(\exists x)(\forall y \in z)$
 $\varphi(x, y, x_1, \dots, x_n) \leftrightarrow (\exists x)(\forall^{\text{st}} y \in A) \varphi(x, y, x_1, \dots, x_n)$
for all $\varphi \in (\text{ZFC})$;
- (2) $(\forall A)(\exists^{\text{st}} {}^*A)(\forall^{\text{st}} x)(x \in A \leftrightarrow x \in {}^*A \wedge \varphi(x))$
for all $\varphi \in (\text{INT})$.

It is worth observing that the Nelson algorithm is mostly operative in INT.

3.4.15. We now describe the theory NST in a version most similar to EXT and IST (in fact Kawai has propounded a somewhat different axiomatics, enabling us to consider the classes of von Neumann–Gödel–Bernays theory as external sets).

3.4.16. The alphabet and conventions of NST are exactly the same as those of EXT. Furthermore, NST accepts all axioms of external set formation, all axioms of interplay between the universes of sets, and the transfer principle of EXT. The differences between NST and EXT lie in the ways of stating the idealization and standardization principles as well in the following supplementary postulate.

3.4.17. Axiom of Acceptance: $\mathbf{V}^{\text{st}} \in \mathbf{V}^{\text{Ext}}$, i.e., the standard universe of NST is an external set.

In view of this axiom, an external set A in NST is of *appropriate size*, in symbols $A \in \mathbf{V}^{\text{a-size}}$, provided there is an external function f from \mathbf{V}^{St} onto A . Note that \mathbf{V}^{St} is of appropriate size. We also agree that the record $a\text{-fin}(A)$ will imply in the sequel that there is a injective external mapping from A onto some standard finite set.

3.4.18. Standardization Principle in NST reads:

$$(\forall A)((\exists^{\text{st}} X) A \subset X \rightarrow (\exists^{\text{st}} {}^*A)(\forall^{\text{st}} x)(x \in A \leftrightarrow x \in {}^*A)).$$

In other words, NST allows only the standardization of the external subsets of standard sets rather than arbitrary external sets as is the case of EXT.

3.4.19. Idealization Principle in NST reads:

$$\begin{aligned} & (\forall^{\text{Int}} x_1) \dots (\forall^{\text{Int}} x_n) (\forall A \in \mathbf{V}^{\text{a-size}}) ((\forall z) z \subset A \wedge a = \text{fin}(z)) \\ & \quad \rightarrow (\exists^{\text{Int}} x) (\forall y \in z) \varphi^{\text{Int}}(x, y, x_1, \dots, x_n) \\ & \quad \rightarrow (\exists^{\text{Int}} x) (\forall^{\text{Int}} y \in A) \varphi^{\text{Int}}(x, y, x_1, \dots, x_n) \end{aligned}$$

for $\varphi \in (\text{ZFC})$.

3.4.20. Kawai Theorem. NST is conservative over ZFC.

◁ The proof proceeds along the lines of the proof of the Powell Theorem on using 3.2.20 (see [232]). ▷

3.4.21. It is worth noting again that the internal universe V^{Int} in NST, furnished with the relativized standardization, idealization, and transfer principles, serves as a model for IST. In other words, we may safely use all technical means of NST for handling the external sets of IST in order to discover the truths of “standard” mathematics.

3.4.22. Taking liberties with notations, we denote the external universe by \mathbf{V}^E (irrespective of whether we imply NST or EXT). Analogously, we let \mathbf{V}^I and \mathbf{V}^S stand for the internal and standard universes. Repeating the scheme of constructing the von Neumann universe, i.e., consecutively taking the unions and powersets of external subsets of already available sets, we start from the empty set and come to the *classical universe* \mathbf{V}^C , the world of “classical sets.” In more detail,

$$\begin{aligned} V_\beta^C & := \{x : (\exists^{\text{st}} \alpha \in \beta)(x \in \mathcal{P}^{\text{Ext}}(V_\alpha^C))\}, \\ \mathbf{V}^C & := \bigcup_{\beta \in \text{On}^{\text{St}}} V_\beta^C, \end{aligned}$$

where On^{St} is the class of standard ordinals. Therefore, the empty set is “classical,” and every “classical” set is composed only of “classical elements.”

3.4.23. Using recursion, i.e., walking about the stores of the classical universe, we may define the robinsonian standardization or $*$ -map.

A standard set $*A$ is the *robinsonian standardization* or the $*$ -image of a “classical” set A provided that every standard element of $*A$ is the $*$ -image of some element of A . In symbols,

$$*\emptyset := \emptyset, \quad *A := *\{*a : a \in A\}.$$

It is worth observing that the soundness of standardization raises no doubts in EXT. The definition of \mathbf{V}^C shows the legitimacy of using this operation in NST to define the robinsonian standardization.

Similar arguments (cf. 3.2.12) demonstrate that the $*$ -map identifies in a bijective manner the universes \mathbf{V}^C and \mathbf{V}^S . Moreover, the robinsonian standardization ensures the validity of the *transfer principle*

$$(\forall x_1 \in \mathbf{V}^C) \dots (\forall x_n \in \mathbf{V}^C) \varphi^C(A_1, \dots, A_n) \leftrightarrow \varphi^S(*A_1, \dots, *A_n)$$

for an arbitrary formula φ of Zermelo–Fraenkel set theory (as usual, φ^C and φ^S are the relativizations of φ to \mathbf{V}^C and \mathbf{V}^S , respectively). It is usual to refer to this form of transfer as the *Leibniz principle*. Since every bijection is usually viewed as identification; therefore, it is a routine practice to denote $*A$ by A while using the $*$ -map. This slightly abuses the language but prevails in common parlance. We will indulge in this sin throughout the sequel.

3.5. Credenda of Infinitesimal Analysis

The discussion of the previous sections has enriched and extended the initial naive views of sets which we use in infinitesimal analysis. From the conventional von Neumann universe \mathbf{V} we came to the internal universe \mathbf{V}^I , the scene of IST presenting in fact the von Neumann universe with reference points, the standard sets forming the standard universe \mathbf{V}^S (Fig. 4).

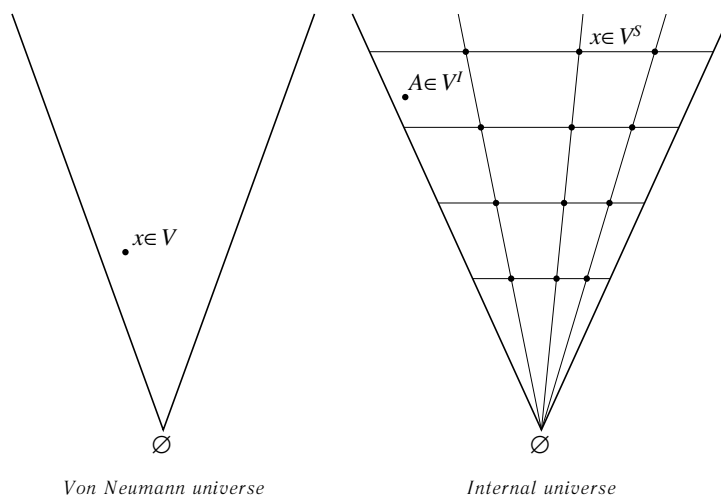


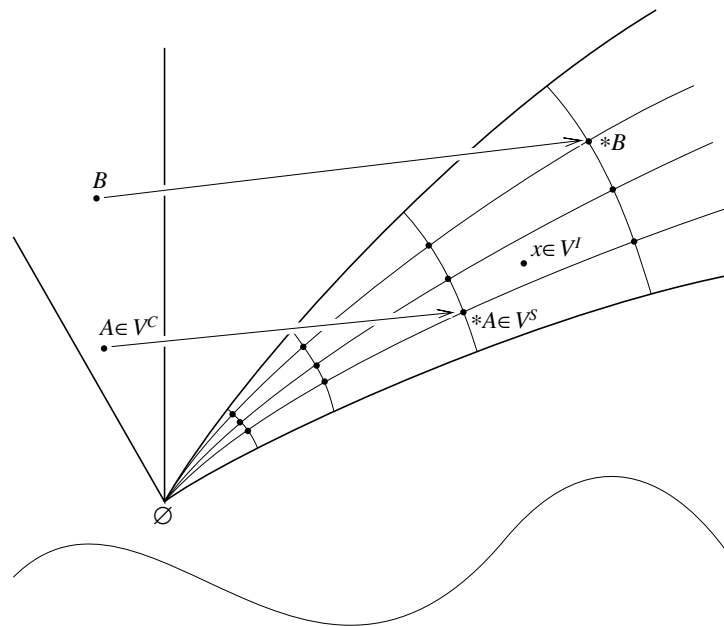
Fig. 4

Further analysis has shown that \mathbf{V}^I lies in a new class \mathbf{V}^E , the external universe collecting all external sets in a Zermelo world. In \mathbf{V}^E we have selected the classical universe \mathbf{V}^C that collects “classical” sets, presenting another implementation of the standard universe. We have also constructed the robinsonian $*$ -map that is an elementwise bijection from \mathbf{V}^C to \mathbf{V}^S . By transfer, we may treat the universes \mathbf{V}^C , \mathbf{V}^S , and \mathbf{V}^I as hypostases of the von Neumann universe (Fig. 5).

3.5.1. This picture of the location \mathbf{V}^E , \mathbf{V}^I , \mathbf{V}^S , and \mathbf{V}^C and the interplay between these worlds of sets drive us to formulating three general set-theoretic stances or credos of infinitesimal analysis. These stances, known as *classical*, *neoclassical*, and *radical*, canonize the general views of the objects and methods of research. The acceptance of one of these stances is essential, determining for example the manner of exposition of the mathematical results obtained by using infinitesimals. Therefore, some familiarity with these stances seems to be an absolute necessity.

3.5.2. The *classical stance* of infinitesimal analysis relates to the technique of its founder Robinson, and the corresponding formalism is mostly common at present.

This credo proclaims that the *principal object of research is the world of classical mathematics* which is identified with the classical universe \mathbf{V}^C . The adept considers \mathbf{V}^C as the “standard universe” (in practice, he or she deals usually with a sufficiently large part of \mathbf{V}^C that contains all particular entities to study, which is the so-called “superstructure”).



External universe

Fig. 5

The main tool for studying the initial “standard universe” is the “nonstandard universe” of internal sets \mathbf{V}^I (or an appropriate part of it) and the $*$ -map that glues together a usual “standard” set and its image in the “nonstandard universe” under the $*$ -map.

It is worth noting that this stance involves a specific usage of the words “standard” and “nonstandard.” The robinsonian standardizations, which are the members of \mathbf{V}^S , are viewed as “nonstandard” objects. A “standard” set is by credo an arbitrary representative of the classical universe \mathbf{V}^C , a member of the “standard” universe. The adept points out that the $*$ -map usually adds new “ideal” elements to sets. He or she implies that $*A = \{ *a : a \in A \}$ if and only if the “classical” or “standard” set A is finite.

For example, if the adept places \mathbb{R} in \mathbf{V}^C and studies the $*$ -image $*\mathbb{R}$ of \mathbb{R} then he or she will see that $*\mathbb{R}$ plays the role of the reals in the sense of the internal universe \mathbf{V}^I . At the same time $*\mathbb{R}$ is not equal to the set of its standard elements: ${}^\circ(*\mathbb{R}) = \{ *t : t \in \mathbb{R} \}$. Considering that $*\mathbb{R}$ is the “internal set of the reals \mathbb{R} ,” and ${}^\circ(*\mathbb{R})$ is its standard part, the adept sometimes takes the liberty of writing ${}^\circ\mathbb{R} := \{ *t : t \in \mathbb{R} \}$ and even $\mathbb{R} := \{ *t : t \in \mathbb{R} \}$. To visualize, the presence of new elements in $*\mathbb{R}$ is expressed as $*\mathbb{R} - \mathbb{R} \neq \emptyset$, and the adept talks of the “hyperreals” $*\mathbb{R}$ extending the “standard” reals \mathbb{R} .

A similar policy is pursued when considering an arbitrary classical set X . Namely, the adept assumes that $X := \{ *x : x \in X \}$ and, therefore, $X \subset *X$. If X is infinite then $*X - X \neq \emptyset$. In other words, the robinsonian standardization adds new elements to all infinite sets. Furthermore, these additional “ideal” entities are galore by idealization in \mathbf{V}^I which is also referred to in this stance as the technique of concurrence or saturation.

3.5.3. Assume that F is a correspondence and A lies in $\text{dom}(F)$. Call F concurrent or finitely satisfiable or directed on A provided that to each nonempty finite subset A_0 of A there is some b satisfying $(a_0, b) \in F$ for all $a_0 \in A_0$.

If the above definition is abstracted by requiring that A_0 has cardinality at most given cardinal κ then we arrive to the definition of κ -concurrence.

3.5.4. Weak Concurrence Principle. To each correspondence F concurrent on A , there is some b in $\text{im}(*F)$ satisfying $(*a, b) \in *F$ for all $a \in A$.

3.5.5. It is easy to see conversely that the validity of 3.5.4 guarantees a natural analog of idealization in a somewhat weaker form than in IST as “relativized to standard sets.” In this regard, applications often involve various conservative enlargements of the classical set theory which allow for concurrence in the form of 3.5.4 and also in stronger forms insuring the additional possibilities of introducing nonstandard entities in a manner that reflects idealization to a full extend.

3.5.6. Strong Concurrence Principle. Let F be such that $*F$ is concurrent on A . Then there is $b \in \text{im}(*F)$ satisfying $(*a, b) \in *F$ for all $a \in A$.

Recall that $(A_\gamma)_{\gamma \in \Gamma}$ is a centered family or has the finite intersection property whenever $\bigcap_{\gamma \in \Gamma_0} A_\gamma \neq \emptyset$ for every nonempty finite subset Γ_0 of Γ .

3.5.7. Saturation Principle. *The following hold:*

- (1) *Let $(A_n)_{n \in \mathbb{N}}$ be a sequence of internal sets enjoying the finite intersection property. Then $\bigcap_{n \in {}^*\mathbb{N}} A_n \neq \emptyset$.*
- (1) *Let $(A_n)_{n \in \mathbb{N}}$ be an increasing sequence of internal sets and $A := \bigcup_{n \in \mathbb{N}} A_n$. Then $A = A_N$ for some $N \in {}^*\mathbb{N}$.*

The definition of ultralimit (cf. 3.2.17) involves arbitrariness which enables us to show that there exist enlargements satisfying the analogs of 3.5.7 and similar principles for each family whose index set has cardinality at most κ . In these circumstances, it is in common parlance to speak of κ -saturation. Rephrased in this terminology, 3.5.7 guarantees ω_0 -saturation. Applications often require ω_1 -saturation as well.

3.5.8. The “enlarged,” “nonstandard” world, i.e., the internal universe \mathbf{V}^I satisfies the transfer principle. Using the properties of the robinsonian standardization, we may rephrase it as follows:

$$(\forall x_1 \in \mathbf{V}^C) \dots (\forall x_n \in \mathbf{V}^C) (\varphi^C(x_1, \dots, x_n) \leftrightarrow \varphi^I({}^*x_1, \dots, {}^*x_n))$$

where φ is an arbitrary formula of Zermelo–Fraenkel set theory. Recall that this form of transfer is the *Leibniz principle*.

3.5.9. Research in the “nonstandard universe” sometimes involves the “*internal set technique*,” i.e., the way of arguing which rests on the fact that every external set given in the “set-theoretic manner” is internal. We now specify one of the versions of this technique.

3.5.10. *Let A be an infinite set. For any set-theoretic formula $\varphi \in (\text{ZFC})$ it is then false that*

$$\{x : \varphi^I(x)\} = {}^*A - A.$$

◁ Assume the contrary. Then the class $\{x : \varphi^I(x)\}$ would be an internal subset of *A . Consequently, A would be internal too. However, the external set ${}^*A - A$ is not internal since A is infinite by hypothesis. ▷

Applications also use other easy versions of the principles of infinitesimal analysis.

3.5.11. *The following hold:*

- (1) **Extension Principle.** *Each sequence $(A_n)_{n \in \mathbb{N}}$ of internal sets A_n is extendible to some internal sequence $(A_n)_{n \in {}^*\mathbb{N}}$;*
- (2) **Overflow Principle.** *If A is an internal set and $\mathbb{N} \subset A$ then A contains some infinitely large hypernatural, i.e. a member of ${}^*\mathbb{N} - \mathbb{N}$;*
- (3) **Underflow Principle.** *If A is an internal set and every infinitely large hypernatural $N \in {}^*\mathbb{N}$ belongs to A then A contains some standard $n \in \mathbb{N}$;*

- (4) **Limitedness Principle.** *If B is an internal set of \mathbb{R} consisting only of limited reals then there is a standard real $t \in \mathbb{R}$ satisfying $B \subset [-t, t]$;*
- (5) **Permanence Principle.** *If B is an internal set containing all positive limited reals then B includes the interval $[0, \Omega]$ for some infinitely large Ω ;*
- (5) **Cauchy Principle.** *If B is an internal set containing all infinitesimals then B includes the interval $[-a, a]$ for some standard $a \in \mathbb{R}$;*
- (6) **Robinson Principle.** *If B is an internal set consisting only of infinitesimals then B is included in the interval $[-\varepsilon, \varepsilon]$ where ε is some infinitesimal.*

It is worth noting that the words “overspill” and “underspill” are in common parlance for “overflow” and underflow.”

3.5.12. Concluding the discussion, we can say that the adept confessing the classical credo works with the two universes, standard and nonstandard. There is a formal possibility of linking the properties of standard and nonstandard objects with the help of the $*$ -map. At the same time the adept may freely translate statements about objects of one universe into those about their images in the other universe, which is the Leibniz principle. The nonstandard universe is abundant in “ideal” elements; various transfinite constructs are realizable in it because of the concurrence principle. The sets falling beyond the nonstandard universe are viewed as external (this is a peculiarity of the terminology: the internal sets are not considered external in this stance). The technique of internal sets proves to be a very effective tool.

The basic advantage of the classical stance is the availability of the $*$ -map enabling us to apply the machinery of infinitesimal analysis to the arbitrary sets of “standard” mathematics. For example, the adept may assert that a function $f : [a, b] \rightarrow \mathbb{R}$ is uniformly continuous if and only if $*f : *[a, b] \rightarrow *\mathbb{R}$ is microcontinuous; i.e., if $*f$ preserves the infinite proximity between the “hyperreals.” The principal complication in absorbing these notions lies in the necessity of imagining the enormous number of the new “ideal” entities inserted forcibly into the ordinary sets. The considerable problems are caused by a natural desire to work (at least in the beginning) with two sets of variables that correspond to the two universes. (When we have constructed the internalization φ^I of a formula φ , we implicitly assume the possibility of this procedure.)

Thus, the *part and parcel of the classical stance, bilingualism and the robinsonian standardization*, determine all its peculiarities, advantages and disadvantages of the corresponding formal machinery.

3.5.13. The **neoclassical stance** of infinitesimal analysis corresponds to the methodology propounded by Nelson. This credo proclaims that the *principal object*

of research is the world of classical mathematics which is identified with the internal universe \mathbf{V}^I included in the external environment \mathbf{V}^E . The “classical” sets do not enter into analysis. The *standard and nonstandard elements are revealed within the ordinary objects*, the internal sets inhabiting \mathbf{V}^I . For instance, the field of reals is the member \mathbb{R} of \mathbf{V}^I which is of course the same as ${}^*\mathbb{R}$, the *hyperreals*, serving as an “ideal” object of the classical stance.

The views we have professed in Chapter 2 correspond mainly to the neoclassical stance. Their advantages are determined by the possibility of studying the “auld lang syne” sets with the goal of discovering new elements in their construction by using additional linguistic means and opportunities. Nelson observed that “really new in nonstandard analysis are not theorems but the notions, i.e., external predicates” [380, p. 134].

The shortcomings of the neoclassical stance stems from the necessity of transferring all definitions and properties from standard objects to their internal relatives by a highly implicit technique of standardization. We have encountered this obstacle before.

3.5.14. The **radical stance** of infinitesimal analysis proclaims that the *principal object of research is the external universe* in full completeness and complexity of its own structure. The adept declares “parochial” or “shy” and discards the classical and neoclassical views of infinitesimal analysis as a technique for study of mathematics basing on Zermelo–Fraenkel set theory.

At a first glance this credo cannot be accepted earnestly and must be dismissed as overextremist. Upon due reflection these accusations of the radical stance should be rejected. This is an illusory, superficial “extremism.” A widely-accepted view of mathematics as the science of forms and relations taken irrespective of their content, as well as the considerably less restrictive classical set-theoretic stance stemming from Cantor, undoubtedly embraces the “extremist” thoughts of the objects of infinitesimal analysis.

Therefore, the most “intrepid” views of sets which we arrive at merge ultimately into the original premises, extending and enriching it. Observe that we started with a “modest” claim that infinitesimal analysis considers as sets exactly the same entities as the rest of mathematics (cf. 2.1.3). We have so come full circle.

3.6. Von Neumann–Gödel–Bernays Theory

As we have already mentioned in 3.2.5, the axiom of replacement ZF_4^φ of Zermelo–Fraenkel set theory ZFC is in fact an axiom-schema embracing infinitely many axioms because of arbitrariness in the choice of a formula φ . It stands to reason to introduce some primitive object that is determined from each formula φ participating in ZF_4^φ . With these objects available, we may paraphrase the content of the axiom-schema ZF_4^φ as a single axiom about new objects. To this end, we need

the axioms that guarantee existence for the objects determined from a set-theoretic formula.

Since all formulas are constructed by a unique procedure in finitely many steps, we find highly plausible the possibility of achieving our goal with finitely many axioms. It is this basic idea stemming from von Neumann that became a cornerstone of the axiomatics of set theory which was elaborated by Gödel and Bernays and is commonly designated by NGB.

The initial undefinable object of NGB is a class. A *set* is a *class* that is a member of some class. A class other than any set is a *proper* class. Objectivization of classes constitutes the basic difference between NGB and ZFC, with the metalanguage of the latter treating “class” and “property” as synonyms.

Axiomatic presentation of NGB uses as a rule one of the two available modifications of the language of ZFC. The first consists in adding a new unary predicate symbol M to the language of ZFC, with $M(X)$ implying semantically that X is a set. The second modification uses two different types of variables for sets and classes. It worth observing that these tricks are not obligatory for describing NGB and reside in exposition routinely for the sake of convenience.

3.6.1. The system NGB is a first-order theory with equality. Strictly speaking, the language of NGB does not differ at all from that of ZFC. However, the upper case Latin letters X, Y, Z, \dots , possible with indices, are commonly used for variables, while the lower case Latin letters are left for the argo resulting from introducing the abbreviations that are absent in the language of NGB.

Let $M(X)$ stand for the formula $(\exists Y)(X \in Y)$. We read $M(X)$ as “ X is a set.”

Introduce the lower case Latin letters x, y, z, \dots (with indices) for the *bound* variables ranging over sets. Speaking more exactly, the formulas $(\forall x)\varphi(x)$ and $(\exists x)\varphi(x)$, called *generalization* and *instantiation* of φ by x , are abbreviations of the formulas $(\forall X)(M(X) \rightarrow \varphi(X))$ and $(\exists X)(M(X) \wedge \varphi(X))$, respectively. Semantically these formulas imply: “ φ holds for every set” and “there is a set for which φ is true.” In this event the variable X must not occur in φ nor in the formulas comprising the above abbreviations.

The rules for using upper case and lower case letters will however be observed only within the present section. On convincing ourselves that the theory of classes may be formalized in principle, we will gradually return to the cozy and liberal realm of common mathematical parlance. For instance, abstracting the set-theoretic concept of function to the new universe of discourse, we customarily speak about a *function-class* F implying that F might be other than a set but still obeys the conventional properties of a function. This is a sacrosanct privilege of the working mathematician.

We now proceed with stating the special axioms of NGB.

3.6.2. Axiom of Extensionality NGB₁. *Two classes coincide if and only if they consist of the same elements:*

$$(\forall X)(\forall Y)(X = Y \leftrightarrow (\forall Z)(Z \in X \leftrightarrow Z \in Y)).$$

3.6.3. We now list the axioms for sets:

(1) Axiom of Pairing NGB₂:

$$(\forall x)(\forall y)(\exists z)(\forall u)(u \in z \leftrightarrow u = x \vee u = y);$$

(2) Axiom of Union NGB₃:

$$(\forall x)(\exists y)(z \in y \leftrightarrow (\exists u)(u \in x \wedge z \in u));$$

(3) Axiom of Powerset NGB₄:

$$(\forall x)(\exists y)(\forall z)(z \in y \leftrightarrow z \subset x);$$

(4) Axiom of Infinity NGB₅:

$$(\exists x)(\emptyset \in x \wedge (\forall y)(y \in x \leftrightarrow y \cup \{y\} \in x)).$$

These axioms coincide obviously with their counterparts in ZFC, cf. A.2.3, A.2.4, A.2.7, and A.2.8. However, we should always bear in mind that the verbal formulations of NGB₁–NGB₅ presume a “set” to be merely a member of another class. Recall also that the lower case Latin letters symbolize abbreviations (cf. 3.6.1). By way of illustration, we remark that, in partially expanded form, the axiom of powerset NGB₄ looks like

$$(\forall X)(M(X) \rightarrow (\exists Y)(M(Y) \wedge (\forall Z)(M(Z) \rightarrow (Z \in Y \leftrightarrow Z \subset X)))).$$

The record of the axiom of infinity uses the following abbreviation

$$\emptyset \in x := (\exists y)(y \in x \wedge (\forall u)(u \notin y)).$$

Existence of the *empty set* is a theorem rather than a postulate in NGB in much the same way as in ZFC. Nevertheless, it is common to enlist the existence of the empty set in NGB as a special axiom:

(5) Axiom of the Empty Set:

$$(\exists y)(\forall u)(u \notin y).$$

3.6.4. Axiom of Replacement NGB₆. *If X is a single-valued class then, for each set y , the class of the second components of those couples of X whose first components belong to y , is a set:*

$$(\forall X)(\text{Un}(X) \rightarrow (\forall y)(\exists z)(\forall u)(u \in z \leftrightarrow (\exists v)((v, u) \in X \wedge v \in y))),$$

where $U_n(X) := (\forall u)(\forall v)(\forall w)((u, v) \in X \wedge (u, w) \in X \rightarrow v = w)$.

As was intended, the axiom-schema of replacement ZF_4^φ turns into a single axiom. Note that the axiom-schema of comprehension in ZF (see 3.2.5) also transforms into a single axiom, the *axiom of comprehension*. The latter reads that, to each set x and each class Y , there is a set consisting of the common members of x and Y :

$$(\forall x)(\forall Y)(\exists z)(\forall u)(u \in z \leftrightarrow u \in x \wedge u \in Y).$$

The axiom of comprehension is weaker than the axiom of replacement since the former ensues from NGB_6 and Theorem 3.6.14 below. However, comprehension is often convenient for practical purposes.

The collection of axioms to follow, NGB_7 – NGB_{13} , relates to the formation of classes. These axioms state that, given some properties expressible by formulas, we may deal with the classes of the sets possessing the requested properties. As usual, uniqueness in these cases results from the axiom of extensionality for classes NGB_1 .

3.6.5. Axiom of Membership NGB₇. *There is a class comprising every couple of sets whose first component is a member of the second:*

$$(\exists X)(\forall y)(\forall z)((y, z) \in X \leftrightarrow y \in z).$$

3.6.6. Axiom of Intersection NGB₈. *There is a class comprising the common members of every two classes:*

$$(\forall X)(\forall Y)(\exists Z)(\forall u)(u \in Z \leftrightarrow u \in X \wedge u \in Y).$$

3.6.7. Axiom of Complement NGB₉. *To each class X there is a class comprising the nonmembers of X :*

$$(\forall X)(\exists Y)(\forall u)(u \in Y \leftrightarrow u \notin X).$$

This implies the existence of the *universal class* $\mathbf{U} := \overline{\emptyset}$ which is the complement of the *empty class* \emptyset .

3.6.8. Axiom of Domain NGB₁₀. To each class X of couples there is a class $Y := \text{dom}(X)$ comprising the first components of the members of X :

$$(\forall X)(\exists Y)(\forall u)(u \in Y \leftrightarrow (\exists v)((u, v) \in X)).$$

3.6.9. Axiom of Product NGB₁₁. To each class X there is a class $Y := X \times \mathbf{U}$ comprising the couples whose first components are members of X :

$$(\forall X)(\exists Y)(\forall u)(\forall v)((u, v) \in Y \leftrightarrow u \in X).$$

3.6.10. Axioms of Permutation NGB₁₂ and NGB₁₃. Assume that $\sigma := (\iota_1, \iota_2, \iota_3)$ is a permutation of $\{1, 2, 3\}$. A class Y is a σ -permutation of a class X provided that $(x_1, x_2, x_3) \in Y$ whenever $(x_{\iota_1}, x_{\iota_2}, x_{\iota_3}) \in X$.

To each class X , there are (2, 3, 1)- and (1, 3, 2)-permutations of X :

$$\begin{aligned} &(\forall X)(\exists Y)(\forall u)(\forall v)(\forall w)((u, v, w) \in Y \leftrightarrow (v, w, u) \in X); \\ &(\forall X)(\exists Y)(\forall u)(\forall v)(\forall w)((u, v, w) \in Y \leftrightarrow (u, w, v) \in X). \end{aligned}$$

The above axioms of class formation proclaim existence of unique classes, as was already observed. It is so in common parlance to speak about *the* complement of a class, *the* intersection of classes, etc.

3.6.11. Axiom of Regularity NGB₁₄. Each nonempty class X has a member having no common elements with X :

$$(\forall X)(X \neq \emptyset \rightarrow (\exists y)(y \in X \wedge y \cap X = \emptyset)).$$

3.6.12. Axiom of Choice NGB₁₅. To each class X there is a choice function-class on X ; i.e., a single-valued class assigning an element of X to each nonempty member of X :

$$(\forall X)(\exists Y)(\forall u)(u \neq \emptyset \wedge u \in X \rightarrow (\exists! v)(v \in u \wedge (u, v) \in Y)).$$

This is a very strong form of the axiom of choice which amounts to a possibility of a simultaneous choice of an element from each nonempty set.

The above axiom makes the list of the special axioms of NGB complete. A moment's inspection shows that NGB, unlike ZFC, has finitely many axioms. Another convenient feature of NGB is the opportunity to treat sets and properties of sets as formal objects, thus implementing the objectivization that is absolutely inaccessible to the expressive means of ZFC.

3.6.13. We now derive a few consequences of the axioms of class formation which are needed in the sequel.

(1) *To each class X there corresponds the (2, 1)-permutation of X :*

$$(\forall X)(\exists Z)(\forall u)(\forall v)((u, v) \in Z \leftrightarrow (v, u) \in X).$$

◁ The axiom of product guarantees existence for the class $X \times \mathbf{U}$. Consecutively applying the axioms of the (2, 3, 1)-permutation and (1, 3, 2)-permutation to the $X \times \mathbf{U}$, arrive at the class Y of 3-tuples (alternatively, *triples*) (v, u, w) such that $(v, u) \in X$. Appealing to the axiom of domain, conclude that $Z := \text{dom}(Y)$ is the sought class. ▷

(2) *To each couple of classes there corresponds their product:*

$$(\forall X)(\forall Y)(\exists Z)(\forall w) \\ (w \in Z \leftrightarrow (\exists u \in X)(\exists v \in Y)(w = (u, v))).$$

◁ To prove the claim, apply consecutively the axiom of product, (1), and the axiom of intersection to arrange $Z := (\mathbf{U} \times Y) \cap (X \times \mathbf{U})$. ▷

Given $n \geq 2$, we may define the class \mathbf{U}^n of all ordered n -tuples by virtue of 3.6.13 (2).

(3) *To each class X there corresponds the class $Z := (\mathbf{U}^n \times \mathbf{U}^m) \cap (X \times \mathbf{U}^m)$:*

$$(\forall X)(\exists Z)(\forall x_1) \dots (\forall x_n)(\forall y_1) \dots (\forall y_m) \\ ((x_1, \dots, x_n, y_1, \dots, y_m) \in Z \leftrightarrow (x_1, \dots, x_n) \in X).$$

(4) *To each class X there corresponds the class $Z := (\mathbf{U}^m \times \mathbf{U}^n) \cap (\mathbf{U}^m \times X)$:*

$$(\forall X)(\exists Z)(\forall x_1) \dots (\forall x_n)(\forall y_1) \dots (\forall y_m) \\ ((y_1, \dots, y_m, x_1, \dots, x_n) \in Z \leftrightarrow (x_1, \dots, x_n) \in X).$$

◁ To demonstrate (3) and (4), apply the axiom of product and the axiom of intersection. ▷

(5) *To each class X there corresponds the class Z satisfying*

$$(\forall x_1) \dots (\forall x_n)(\forall y_1) \dots (\forall y_m) \\ ((x_1, \dots, x_{n-1}, y_1, \dots, y_m, x_n) \in Z \leftrightarrow (x_1, \dots, x_n) \in X).$$

◁ Appeal to the axioms of permutation and the axiom of product. ▷

3.6.14. Theorem. Let φ be a formula whose variables are among $X_1, \dots, X_n, Y_1, \dots, Y_m$ and which is predicative; i.e., all bound variables of φ range over sets. Then the following is provable in NGB:

$$(\forall Y_1) \dots (\forall Y_m) (\exists Z) (\forall x_1) \dots (\forall x_n) \\ ((x_1, \dots, x_n) \in Z \leftrightarrow \varphi(x_1, \dots, x_n, Y_1, \dots, Y_m)).$$

◁ Assume that φ is written so that the only bound variables of φ are those for sets. It suffices to consider only φ containing no subformulas of the shape $Y \in W$ and $X \in X$, since these subformulas might be rewritten in equivalent form as $(\exists x)(x = Y \wedge x \in W)$ and $(\exists u)(u = X \wedge u \in X)$. Moreover, the symbol of equality may be eliminated from φ on substituting for $X = Y$ the expression $(\forall u)(u \in X \leftrightarrow u \in Y)$, which is sound by the axiom of extensionality. The proof proceeds by induction on the *complexity* or *length* k of φ ; i.e., by the number k of propositional connectives and quantifiers occurring in φ .

In case $k = 0$ the formula φ is atomic and has the form $x_i \in x_j$, or $x_j \in x_i$, or $x_i \in Y_l$ ($i < j \leq n, l \leq m$). If $\varphi := x_i \in x_j$ then, by the axiom of membership, there is a class W_1 satisfying

$$(\forall x_i)(\forall x_j)((x_i, x_j) \in W_1 \leftrightarrow x_i \in x_j).$$

If $\varphi := x_j \in x_i$ then, using the axiom of membership again, we find a class W_2 with the property

$$(\forall x_i)(\forall x_j)((x_j, x_i) \in W_2 \leftrightarrow x_j \in x_i),$$

and apply 3.6.13 (1). In result, we obtain a class W_3 such that

$$(\forall x_i)(\forall x_j)((x_i, x_j) \in W_3 \leftrightarrow x_j \in x_i).$$

Hence, in each of these two cases there is a class W satisfying the following formula:

$$\Phi := (\forall x_i)(\forall x_j)((x_i, x_j) \in W \leftrightarrow \varphi(x_1, \dots, x_n, Y_1, \dots, Y_m)).$$

By 3.6.13 (4), we may replace the subformula $(x_i, x_j) \in W$ of Φ with the membership $(x_1, \dots, x_{i-1}, x_i) \in Z_1$ for some other class Z_1 and insert the quantifiers $(\forall x_1) \dots (\forall x_{i-1})$ in the prefix of Φ .

Let Ψ be the so-obtained formula. By 3.6.13 (5), there is some class Z_2 for Ψ so that it is possible to write $(x_1, \dots, x_{i-1}, x_i, x_j) \in Z_1$ instead of the subformula $(x_1, \dots, x_i, x_{i+1}, \dots, x_j) \in Z_2$ and to insert the quantifiers $(\forall x_{i+1}) \dots (\forall x_{j-1})$ in the prefix of Ψ . Finally, on applying 3.6.13 (3) to Z_2 , find a class Z satisfying the following formula:

$$(\forall x_1) \dots (\forall x_n)((x_1, \dots, x_n) \in Z \leftrightarrow \varphi(x_1, \dots, x_n, Y_1, \dots, Y_m)).$$

In the remaining case of $x_i \in Y_l$, the claim follows from existence of the products $W := \mathbf{U}^{i-1} \times Y_l$ and $Z := W \times \mathbf{U}^{n-i}$. This completes the proof of the theorem for $k = 0$.

Assume now that the claim of the theorem is demonstrated for all $k < p$ and the formula φ has p propositional connectives and quantifiers. It suffices to consider the cases in which φ results from some other formulas by negation, implication, and generalization.

Suppose that $\varphi := \neg\psi$. By the induction hypothesis, there is a class V such that

$$(\forall x_1) \dots (\forall x_n)((x_1, \dots, x_n) \in V \leftrightarrow \psi(x_1, \dots, x_n, Y_1, \dots, Y_m)).$$

By the axiom of complement, the class $Z := \mathbf{U} - V := \mathbf{U} \setminus V$ meets the required conditions.

Suppose that $\varphi := \psi \rightarrow \theta$. Again, by the induction hypothesis, there are classes V and W making the claim holding for V and ψ and such that

$$(\forall x_1) \dots (\forall x_n)((x_1, \dots, x_n) \in W \leftrightarrow \theta(x_1, \dots, x_n, Y_1, \dots, Y_m)).$$

The sought class $Z := \mathbf{U} - (V \cap (\mathbf{U} - W))$ exists by the axioms of intersection and complement.

Suppose that $\varphi := (\forall x)\psi$, and let V and ψ be the same as above. Applying the axiom of domain to the class $X := \mathbf{U} - V$, obtain the class Z_1 such that

$$(\forall x_1) \dots (\forall x_n)((x_1, \dots, x_n) \in Z_1 \leftrightarrow (\exists x)\neg\psi(x_1, \dots, x_n, Y_1, \dots, Y_m)).$$

The class $Z := \mathbf{U} - Z_1$ exists by the axiom of complement and is the one we seek since the formula $(\forall x)\psi$ amounts to $\neg(\exists x)(\neg\psi)$. \triangleright

3.6.15. Each of the axioms of class formation NGB₇–NGB₁₃ is a corollary to Theorem 3.6.14 provided that the formula φ is duly chosen. On the other hand, the theorem itself, as shown by inspection of its proof, ensues from the axioms of class formation. Remarkably, we are done on using finitely many axioms NGB₇–NGB₁₃ rather than infinitely many assertions of Theorem 3.6.14.

Theorem 3.6.14 allows us to prove the existence of various classes. For instance, to each class Y there corresponds the class $\mathcal{P}(Y)$ of all subsets of Y , as well as the union $\bigcup Y$ of all elements of Y . These two classes are defined by the conventional formulas:

$$\begin{aligned} (\forall u)(u \in \mathcal{P}(Y) &\leftrightarrow u \subset Y), \\ (\forall u)(u \in \bigcup Y &\leftrightarrow (\exists v)(v \in Y \wedge u \in v)). \end{aligned}$$

The above claims of existence are easy on putting $\varphi(X, Y) := X \subset Y$ and $\varphi(X, Y) := (\exists V)(X \in V \wedge V \in Y)$. Analogous arguments corroborate the definitions of Z^{-1} , $\text{im}(Z)$, $Z \upharpoonright Y$, $Z \text{“} Y$, $X \cup Y$, etc., with X , Y , and Z arbitrary classes.

3.6.16. Theorem. *Each theorem of ZFC is a theorem of NGB.*

◁ Each axiom of ZFC is a theorem of NGB. The only nonobvious part of the claim concerns the axiom of replacement ZF_4^φ which we will prove.

Assume that y is not unbound in φ , and let $\{x, t, z_1, \dots, z_m\}$ stand for the complete list of variables in the construction of φ . Assume further that, for all x, u, v, z_1, \dots, z_m , the following formula holds:

$$\varphi(x, u, z_1, \dots, z_m) \wedge \varphi(x, v, z_1, \dots, z_m) \rightarrow u = v.$$

The formula is predicative since each bound variable of φ ranges over sets. By Theorem 3.6.14, there is a class Z such that

$$(\forall x)(\forall u)((x, u) \in Z \leftrightarrow \varphi(x, u, z_1, \dots, z_m)).$$

This property of φ shows that the class Z is single-valued; i.e., $\text{Un}(Z)$ is provable in NGB. By the axiom of replacement NGB_6 , there is a set y satisfying

$$(\forall v)(v \in y \leftrightarrow (\exists u)((u, v) \in Z \wedge u \in x)).$$

Obviously, y satisfies the desired formula

$$(\forall z_1) \dots (\forall z_m)(\forall v)(v \in y \leftrightarrow (\exists u \in x) \varphi(u, v, z_1, \dots, z_m)). \triangleright$$

3.6.17. Theorem. *Each formula of ZFC expressing a theorem of NGB is a theorem of ZFC.*

◁ The proof may be found, for instance, in [62]. It uses some general facts of model theory which lie beyond the framework of the present book. ▷

Theorems 3.6.16 and 3.6.17 are often paraphrased as follows.

3.6.18. Theorem. *Von Neumann–Gödel–Bernays set theory is conservative over Zermelo–Fraenkel set theory.*

3.6.19. Among the other axiomatic set theories, we mention the so-called Bernays–Morse theory that extends NGB. Bernays–Morse set theory assumes the special axioms NGB_1 – NGB_5 , NGB_{14} and the following axiom-schema of comprehension:

$$(\exists X)(\forall Y)(Y \in X \leftrightarrow M(Y) \wedge \varphi(Y, X_1, \dots, X_n)),$$

with φ an arbitrary formula without free occurrences of X .

3.7. Nonstandard Class Theory

In this section we introduce NCT, another axiomatics analogous to IST but differing from IST since NCT is an extension of von Neumann–Gödel–Bernays class theory. Each of the principles of transfer, idealization, and standardization is stated in NCT as an axiom rather than an axiom-schema in IST, and so NCT, as well as NGB, is finitely axiomatizable.

3.7.1. The language of NCT results from the language of NGB by supplementing the alphabet of NGB with a new symbol St for the unary predicate expressing the property of a class to be or not to be standard (the record $\text{St}(X)$ reads: “ X is a standard class”). In much the same way as in 3.6, the variables ranging over classes are denoted by capital Latin letters while the variable ranging over sets are denoted by lowercase Latin letters.

We will use other abbreviations and conventions of 3.6. In particular, a class X is a *set* in symbols $M(X)$, provided that X is a member of some class: $M(X) := (\exists Y)(X \in Y)$ (cf. 3.6.1). As before, the record $S(X_1, \dots, X_n) := \varphi(X_1, \dots, X_n)$ means that $S(X_1, \dots, X_n)$ is an abbreviation of $\varphi(X_1, \dots, X_n)$.

All in all, the language of NCT is the language of a first order predicate calculus with equality which contains one unary predicate symbol \in and one unary predicate symbol St . We now list the special axioms of NCT.

(1) In NCT we accept the same axioms of extensionality, pairing, union, powersets, infinity, and choice as in NGB, namely, NBG_1 – NBG_5 , NBG_{14} , NBG_{15} (see 3.6.2, 3.6.3, 3.6.11, and 3.6.12).

(2) **Axiom of Replacement** in NCT reads:

$$(\forall V)(\forall x)(\exists y)(\forall u \in x)(\exists v((u, v) \in V) \rightarrow (\exists v \in y)((u, v) \in V)).$$

In the sequel we will use the following abbreviations:

$$\begin{aligned} (\exists^{\text{st}}x)\varphi &:= (\exists x)(\text{St}(x) \wedge \varphi), \\ (\forall^{\text{st}}x)\varphi &:= (\forall x)(\text{St}(x) \rightarrow \varphi). \end{aligned}$$

(3) **Axiom of Boundedness:**

$$(\forall x)(\exists^{\text{st}}z)(x \in z).$$

(4) **Axiom of Transfer:**

$$(\forall^{\text{st}}X)((\exists x)x \in X \rightarrow (\exists^{\text{st}}x)x \in X).$$

(5) **Axiom of Standardization:**

$$(\forall X)(\exists^{\text{st}}Y)(\forall^{\text{st}}x)(x \in Y \leftrightarrow x \in X).$$

Starting with the empty set and appealing to the axiom of standardization, we arrive at the standard class L without any standard elements. By transfer $L = \emptyset$; i.e., *the empty set is standard*.

3.7.2. A formula of NCT is *predicative*, provided that each bound variable ranges over members of some set and the standardness predicate enters only in external quantifiers, i.e., all occurrences of quantifiers and the standardness predicate look like $\exists x$, $\exists^{\text{st}}x$, $\forall x$, and $\forall^{\text{st}}x$. Observe that we may replace the subformula $\text{St}(x)$ with $(\exists^{\text{st}}y)(y = x)$.

Let p be an arbitrary set. A class X is *p -standard* (in symbols, $\text{st}_p X$) provided that X is the *p -section* of some standard class Y ; i.e., $(\exists^{\text{st}}Y)(X = Y''p)$ with $Y''p = \{v : (p, v) \in Y\}$. A class X is *internal*, (in symbols, $\text{int } X$) provided that X is p -standard for some p .

Axiom-Schema of Class Formation: *If $\varphi(x_1, \dots, x_n, Y_1, \dots, Y_m)$ is a predicative formula then*

- (1) *To all classes Y_1, \dots, Y_m there corresponds the class*

$$\mathcal{T} = \{(x_1, \dots, x_n) : \varphi(x_1, \dots, x_n, Y_1, \dots, Y_m)\};$$
- (2) *If φ is an internal formula and Y_1, \dots, Y_m are standard classes then so is \mathcal{T} .*

In much the same way as in NGB, we may accept not the above axiom-schema of class formation but rather finitely many of its particular cases (cf. NGB₇–NGB₁₃) implying the validity of the full version of the axiom-schema (see 3.6.14). Consequently, NCT is a finitely axiomatizable theory.

The following proposition is immediate from the axiom-schema of class formation.

3.7.3. *If φ and Y_1, \dots, Y_m are internal in the axiom-schema of class formation then \mathcal{T} is an internal class. Moreover, if all Y_i are p -standard for the same p then \mathcal{T} is a p -standard class.*

◁ Everything follows from the axiom-schema of class formation. ▷

We now agree to use the notations:

$$\begin{aligned} \mathbb{U} &:= \{x : x = x\} = \{x : x \notin \emptyset\}, \\ \mathbb{E} &:= \{x : (\exists u)(\exists v)(x = (u, v) \wedge u \in v)\}, \\ \mathbb{S} &:= \{x : \text{St}(x)\}, \\ {}^-X &:= \{x : x \notin X\}. \end{aligned}$$

By the axiom-schema of class formation, \mathbb{U} and \mathbb{E} are standard classes, \mathbb{S} is a class. Moreover, if X and Y are some classes then ${}^-X$, $X \cap Y$, $\text{dom}(X)$, and $X \times \mathbb{U}$ are classes which are standard whenever so are X and Y .

Every set x is x -standard and so x is internal since $x = \mathbb{E}^{-1}x$. Every standard class X is internal since $X = (\{\emptyset\} \times X)''\emptyset$.

3.7.4. The following two axioms deal with some properties of internal classes.

(1) **Axiom of Comprehension:**

$$(\forall^{\text{int}} X)(\forall x)(\exists y)(\forall u)(u \in y \leftrightarrow u \in x \wedge u \in X).$$

(2) **Axiom of Idealization:**

$$\begin{aligned} & (\forall^{\text{int}} X)(\forall^{\text{st}} a_0)((\forall^{\text{st fin}} c \subseteq a_0)(\exists x)(\forall a \in c)((x, a) \in X) \\ & \leftrightarrow (\exists x)(\forall^{\text{st}} a \in a_0)((x, a) \in X)). \end{aligned}$$

3.7.5. The following propositions are straightforward from the axiom-schema of class formation and 3.7.3.

(1) *Let φ be an internal predicative formula. Then*

$$\begin{aligned} & (\forall^{\text{int}} X_1) \dots (\forall^{\text{int}} X_n)(\forall x)(\exists y)(\forall u)(u \in y \\ & \leftrightarrow y \in x \wedge \varphi(x, X_1, \dots, X_n)). \end{aligned}$$

In particular, the axiom-schema of comprehension of BST is valid.

(2) *The standardization and bounded idealization principles of BST hold in NST.*

(3) **Transfer Principle.** *If φ is an internal predicative formula then*

$$\begin{aligned} & (\forall^{\text{st}} X_1) \dots (\forall^{\text{st}} X_n)((\forall^{\text{st}} x) \varphi(x, X_1, \dots, X_n) \\ & \rightarrow (\forall x) \varphi(x, X_1, \dots, X_n)). \end{aligned}$$

In particular, the transfer principle of BST is valid.

(4) *Each sentence provable in BST is provable in NCT.*

We emphasize that the axiom-schemata of transfer, idealization, standardization, and comprehension of BST are particular instances of the analogous axioms of NCT in which classes are determined by predicative formulas with free set variables (these formula are internal for the axioms of comprehension, transfer, and idealization).

3.7.6. We note the following propositions:

(1) *If x and p are sets then*

$$\text{st}_p x \leftrightarrow (\exists^{\text{st}} z)(x = z''p) \leftrightarrow (\exists^{\text{st}} f)(\text{Fnc } f \wedge x = f(p)).$$

\triangleleft Assume first that x is a p -standard set. Then $x = z''p$ for some standard z by the axioms of boundedness and transfer. Therefore, $f = \{(q, z''q) : q \in \text{dom}(z)\}$ is a standard function and $f(p) = x$.

Conversely, if f is a standard function then, by the axiom of transfer, $f(p)$ is the p -section of the standard set $\{(q, u) : u \in f(q)\}$. \triangleright

(2) Let φ be an internal predicative formula, and let p be a set. Then

$$\begin{aligned} & (\forall_p^{\text{st}} X_1) \dots (\forall_p^{\text{st}} X_n) ((\forall_p^{\text{st}} x) \varphi(x, X_1, \dots, X_n)) \\ & \rightarrow (\forall x) \varphi(x, X_1, \dots, X_n). \end{aligned}$$

◁ By 3.7.3 it suffices to prove that each nonempty p -standard class X contains a p -standard element. Assume that $X = Y''p$, with $\text{st} Y$ and $p \in r$ for some standard r . By the axioms of comprehension, choice, and transfer there is a standard function f satisfying

$$(\forall q \in r)((\exists y)(q, y) \in Y \rightarrow (\exists y)(q, y) \in Y \cap f).$$

Since X is nonempty; therefore, $p \in \text{dom}(f)$ and $f(p)$ is a p -standard member of X . ▷

3.7.7. Given an arbitrary class C , put ${}^\circ C := C \cap \mathbb{S}$. The axiom of standardization postulates that to each class X there corresponds a standard class Y such that ${}^\circ Y = {}^\circ X$. By transfer, such a standard class is unique. We denote it by ${}^s X$.

(1) **Theorem.** *A class X is standard if and only the intersection of X with each standard set is a standard set.*

◁ Necessity follows from the axiom-schema of class formation. To prove sufficiency, assume that $(\forall^{\text{st}} z)(\exists^{\text{st}} t)(t = X \cap z)$. Put $Y := {}^s\{x : x \in X\}$ and demonstrate that $Y = X$. By the axiom of boundedness it suffices to check that $z \cap X = z \cap Y$ for every standard set z . From the choice of Y it follows that

$${}^\circ(X \cap z) = {}^\circ X \cap {}^\circ z = {}^\circ Y \cap {}^\circ z = {}^\circ(Y \cap z).$$

Since $X \cap z$ and $Y \cap z$ are standard sets, the last chain of equalities implies the claim by transfer. ▷

(2) *A set is standard and finite if and only if its every element is standard.*

◁ By idealization

$$\begin{aligned} \text{St}(x) \wedge \text{fin}(x) & \leftrightarrow (\exists^{\text{st}} \text{fin} y \subseteq x)(\forall a \in x)(\exists b \in y)(a = b) \\ & \leftrightarrow (\forall a \in x)(\exists^{\text{st}} b \in x)(a = b), \end{aligned}$$

which completes the proof. ▷

We call a set *standardly finite* whenever its cardinality is a standard natural.

3.7.8. Theorem. *A set is standardly finite if and only if its every subclass is a set.*

◁ Assume that x is a set, $|x| = \alpha$, and $f : \alpha \rightarrow x$ is a bijection.

If x is not a standardly finite set then $(\forall^{\text{st}} n \in \mathbb{N})(\alpha > n)$ by transfer. Considering the class $I := \{f(n) : n \in {}^\circ\mathbb{N}\}$, we may write

$$(\forall^{\text{st}} \text{fin } s \subseteq \mathbb{N})(\exists k \in \mathbb{N})(\forall n \in s)(f(k) \in I \wedge n < k).$$

Were I set, there would exist some $k \in \mathbb{N}$ such that $f(k) \in I$ and $(\forall^{\text{st}} n \in \mathbb{N})(n < k)$, which is impossible. Indeed, $f(k) \in I$ implies that k is standard since f is a bijection.

Conversely, let x be standardly finite and $X \subseteq x$. Consider the class $T := \{n \in \alpha : f(n) \in X\}$. By standardization and transfer, there is some set t such that $t = {}^s T$ and $t \subseteq \alpha$. Since $\alpha \subseteq \mathbb{S}$ by 3.7.7(2); therefore, $t = {}^\circ t = {}^\circ T = T$. Consequently, $X = \{f(n) : n \in t\}$ is a set. ▷

3.7.9. The p -*monad* of a set x (in symbols $\mu_p(x)$) is the intersection of all p -standard classes including x . Since the complement of a standard class is a standard class too, the p -monads of two sets are coincident or disjoint. From 3.7.6(1) it follows that

$$\mu_p(x) = \{y : (\forall_p^{\text{st}} z)(y \in z \leftrightarrow x \in z)\}.$$

If x is p -standard then we call the class $\mu_p(x)$ the *monad* of x and denote it by $\mu(x)$. Clearly, $\mu(x) = \bigcap \{a \in \mathbb{S} : x \in a\}$.

Look at an arbitrary set x . By the axiom of boundedness, $x \in x_0$ for some standard x_0 . By transfer, it is easy to prove that $u := {}^s \{a \subseteq x_0 : x \in a\}$ is a standard ultrafilter and $\bigcap {}^\circ u = \mu(x)$. Conversely, if u is an arbitrary standard ultrafilter then, by transfer and idealization, $\bigcap {}^\circ u \neq \emptyset$ and $\mu(x) = \bigcap {}^\circ u$ for every $x \in \bigcap {}^\circ u$.

The class $\bigcap {}^\circ u$ is the *nest* of an ultrafilter u ; in symbols, $\nu(u)$. The class of all ultrafilters will be denoted by Ult , while letting $\text{Ult}(x)$ stand for the set of all ultrafilters on a set x .

If x and p are sets then $\mu_p(x) = \mu((p, x))''p$.

◁ From 3.7.6(1) it follows that

$$\begin{aligned} y \in \mu((p, x))''p &\leftrightarrow (p, y) \in \mu((p, x)) \\ &\leftrightarrow (\forall^{\text{st}} z)((p, x) \in z \leftrightarrow (p, y) \in z) \leftrightarrow y \in \mu_p(x), \end{aligned}$$

which completes the proof. ▷

3.7.10. A class X is p -saturated provided that X contains the p -monad of its every element.

(1) A set x is p -standard if and only if x is p -saturated.

◁ Suppose that x is p -saturated. Take an arbitrary $u \in x$ and show that u belongs to the p -section of some standard set lying in x . Indeed, assuming the contrary, note that

$$(\forall^{\text{st}}z)(u \in z''p \rightarrow (\exists v \in z''p)(v \notin x)).$$

We may restrict the range of z in this formula to the standard set $\{t : t \subseteq \bigcup x_0\}$, where x_0 is standard and $x \in x_0$. By idealization,

$$(\exists v \notin x)(\forall^{\text{st}}z)(u \in z''p \rightarrow v \in z''p),$$

which contradicts the inclusion $\mu_p(u) \subseteq x$.

Therefore, $(\forall u \in x)(\exists^{\text{st}}z)(u \in z''p \subseteq x)$. Again by idealization, find a standard finite set z_0 satisfying $(\forall u \in x)(\exists z \in z_0)(u \in z''p \subseteq x)$. It is easy that x serves as the p -section of the standard set $\bigcup z_0$. ▷

(2) $\mu_p(x) = \{x\} \leftrightarrow \text{st}_p x$ for arbitrary sets x and p .

◁ The implication to the left is obvious. If, conversely, $\mu_p(x) = \{x\}$ then $\{x\}$ is p -saturated and, hence, p -standard. By transfer, x is p -standard too. ▷

3.7.11. Axiom of Saturation:

$$(\forall X)(\exists p)(\forall x \in X)(\mu_p(x) \subseteq X),$$

i.e. each class is p -saturated for some set p .

Given a class $D \subseteq \text{Ult}$ and a set p , put

$$\text{Psls}(D, p) := \bigcup_{u \in {}^\circ D} \nu(u)''p.$$

A *semiset* is a subclass of a set:

$$\text{Sms } X := (\exists^{\text{st}}z)(X \subseteq z).$$

3.7.12. Theorem. Let X be an arbitrary class. Then there are some standard class $D \subseteq \text{Ult}$ and a set p such that

$$X = \text{Psls}(D, p).$$

Moreover, if X is a semiset then D may be chosen to be a set.

◁ Let X be a p -saturated class. Put $D = {}^s\{u \in \text{Ult} : \nu(u) \upharpoonright p \subseteq X\}$. Then the equality in question follows from the proposition in 3.7.9. The same equality holds in case $X \in z$, with z standard, provided that as D we take $d := D \cap \text{Ult}(r \times z)$, where r is some standard set containing p . Observe that d is standard by transfer. ▷

Thus, each semiset in NCT turns out to be definable by some predicative Σ_2^{st} -formula. It is possible to prove also that if all quantifiers of a formula φ are restricted to semisets then φ is equivalent to some predicative formula. Indeed, we are done on replacing all subformulas like $\text{st } X$ with $(\forall^{\text{st}} s)(\exists^{\text{st}} t)(t = X \cap s)$ and the formulas like $(\exists X)(\text{Sms } X \rightarrow \varphi(X, \dots))$ with $(\exists^{\text{st}} d)(\exists p) \varphi(\text{Psls}(d, p), \dots)$.

The following theorem is the saturation principle in its traditional statement. Observe that, in contrast to NCT, this theorem cannot be proven or even formulated in IST or BST.

3.7.13. Theorem. *Let a class X and a standard set z_0 be such that*

$$(\forall^{\text{st}} x \in z_0)(\exists y)((x, y) \in X).$$

Then there is a function-set f satisfying

$$(\forall^{\text{st}} x \in z_0)((x, f(x)) \in X).$$

◁ By the axioms of comprehension and boundedness there is a standard set t such that $(\forall x \in z_0)((\exists y)(x, y) \in X \rightarrow (\exists y \in t)(x, y) \in X)$. Assume that X is a p -saturated class. If $(x, y) \in X$ and x is standard, then $(\forall y' \in \mu(y))(x, y') \in X$ since $\mu_p((x, y)) = \{x\} \times \mu_p(y)$. Put $d := {}^s\{(x, u) \in z \times \text{Ult}(t) : \{x\} \times (\nu(u) \upharpoonright p) \subseteq X\}$. The axioms of choice and transfer enable us to find a standard function $h : z_0 \rightarrow \text{Ult}(t)$ satisfying $(\forall x \in z_0)((x, h(x)) \in d)$. Hence,

$$(\forall^{\text{st fin}} z \in z_0)(\exists f)(\forall x \in z)(\forall^{\text{st}} a \in h(x))(\text{Fnc}(f) \wedge f(x) \in a).$$

By idealization, there is a function f such that $(\forall^{\text{st}} x \in z)(f(x) \in \nu(h(x)))$. Obviously, this f completes the proof. ▷

3.7.14. COMMENTS.

(1) The nonstandard class theory NCT, we set forth in this section, was suggested by Andreev and Gordon in [14]. NCT is remarkable for clarity and simplicity. In particular, NCT contains only finitely many axioms and the principles of infinitesimal analysis become axioms rather than axiom-schemata in IST.

(2) The presence of classes enables us to implement various constructions with external sets, which is impossible in IST. In particular, the list of axioms of NCT contains the axiom of saturation (cf. 3.7.11) which plays a key role in applications of infinitesimal analysis.

(3) All sets in NCT are internal. The external objects of NCT are proper classes. Furthermore, in much the same way as alternative set theory AST by Vopěnka [513], NCT allows for subclasses of sets which might fail to be sets (the axiom of comprehension holds only for internal sets). We call these semisets as minted by Vopěnka. The theory NCT enjoys some other properties of AST. In particular, it is a theorem of NCT that x is a standardly finite set if and only if x includes no semisets.

(4) If some internal class (in particular, internal set) X is the section of a standard class by a set p then X is standard relative to p or p -standard, see 3.7.2. This concept of relative standardness in IST was first suggested in the article [141] which showed for instance that the transfer principle and the implication to the right in the idealization principle remain valid on replacing all occurrences of the standardness predicate in them with the predicate of standardness relative to an arbitrary set p . This holds true in NCT (see 3.7.7 (1)).

3.8. Consistency of NCT

In this section we prove that NCT is conservative over BST.

3.8.1. Theorem. *Every predicative proposition provable in NCT is provable in BST.*

◁ We will demonstrate below that each model of BST embeds isomorphically into some model of NCT as the universe of all sets, which implies the claim by the celebrated Gödel Completeness Theorem. ▷

3.8.2. Consider an arbitrary model $\mathfrak{M} = (M, \in^M, \text{st}^M)$ of BST. Let L stand for the enrichment of the language of BST with the elements of M viewed as new symbols. Assume that \mathfrak{M} is a model of the language L on defining the interpretation of a symbol a in M to be a itself. The sets in M , entering a formula φ of the language L , are referred to as the parameters of φ .

Given a formula φ of L with a single unbound variable, put $[\varphi] := \{x : \mathfrak{M} \models \varphi(x)\}$. Furthermore,

$$N := \{[\varphi] : \varphi \text{ is a formula of } L \text{ with a single unbound variable}\};$$

$$\text{Std} := \{[\varphi] \in N : \varphi \text{ is an internal formula with unbound parameters}\};$$

$$\text{Set}(a) := [x \in a] \text{ for all } a \in M.$$

Given $p, q \in N$, define

$$p \in^N q := (\exists a \in M)(p = \text{Set}(a) \wedge a \in q),$$

$$\text{st}^N p := p \in \text{Std}.$$

3.8.3. Assume that $a, b \in M$ and $p, q \in N$. Then

- (1) $p \in^N q \rightarrow (\exists a \in M)(p = \text{Set}(a))$;
- (2) $\text{Set}(a) = \text{Set}(b) \leftrightarrow a = b$;
- (3) If $p = \text{Set}(a)$ and $q = \text{Set}(b)$ then $p \in^N q \leftrightarrow a \in^M b$;
- (4) If $p = \text{Set}(a)$ then $\text{st}^N p \leftrightarrow \text{st}^M a$.

◁ (1): Immediate from the definition of \in^N .

(2): This follows from the validity of the axiom of extensionality in \mathfrak{M} .

(3): This ensures from (1) by the definition of \in^N .

(4): Note that from the definition of st^N it follows that

$$p = \{b : \mathfrak{M} \models b \in a\} = \{b : \mathfrak{M} \models \varphi(b)\},$$

where φ is an internal formula with standard parameters. Consequently, $\mathfrak{M} \models (\forall x)(x \in a \leftrightarrow \varphi(x))$. The axiom-schema of transfer is satisfied in \mathfrak{M} , implying that $\mathfrak{M} \models \text{st} a$, i.e., $\text{st}^M a$.

Conversely, if $\text{st}^M a$ then $p = [x \in a] \in \text{Std}$. ▷

- (5) The mapping Set embeds isomorphically \mathfrak{M} as a model of L into the model $\mathfrak{N} = (N, \in^N, \text{st}^N)$; moreover, $\mathfrak{N} \models (\exists X)(p \in X) \rightarrow (\exists a \in M)(p = \text{Set}(a))$ for all $p \in N$.

◁ Everything follows from (1)–(4). ▷

Proposition 3.8.3 (5) shows that a class p is a set in \mathfrak{N} if and only if $p = \text{Set}(a)$ for some $a \in M$; i.e., \mathfrak{M} does embed into \mathfrak{N} as the universe of all sets.

3.8.4. We are left with checking the validity of the axioms of NCT in the model \mathfrak{M} .

From 3.8.3 (5) it follows that the axioms of NCT, written as predicative formulas, hold in \mathfrak{N} whenever they are valid in BST. This concerns the axioms of pairing, union, powersets, infinity, choice, regularity, and boundedness. The axiom of extensionality holds in \mathfrak{N} by the construction of \in^N .

If φ is a formula of the language L then we let the symbol \mathbf{C}_φ stand for the collection $\{x : \varphi(x, x_1, \dots, x_n)\}$. Assume that $\Phi(X_1, \dots, X_n)$ is a predicative formula, and $\varphi_1(x, x_1, \dots, x_m), \dots, \varphi_n(x, x_1, \dots, x_m)$ are some formulas of L whose unbound variables do not enter the construction of Φ . Denote by $\Phi(\mathbf{C}_{\varphi_1}, \dots, \mathbf{C}_{\varphi_n})$ the formula that results from $\Phi(X_1, \dots, X_n)$ by replacing

- (1) all occurrences of the atomic formulas like $y \in X_j$ with the formula $\varphi_j(y, x_1, \dots, x_m)$;
- (2) all occurrences of the atomic formulas like $X_i \in X_j$ with

$$(\exists x)((\forall y)(y \in x \leftrightarrow \varphi_i(y, x_1, \dots, x_m)) \wedge \varphi_j(x, x_1, \dots, x_m));$$

(3) all occurrences of the atomic formulas like $X_i = X_j$ with

$$(\forall x)(\varphi_i(x, x_1, \dots, x_m) \leftrightarrow \varphi_j(x, x_1, \dots, x_m));$$

(4) all occurrences of the atomic formulas like $X_i = x$ with

$$(\forall y)(y \in x \leftrightarrow \varphi_i(y, x_1, \dots, x_m)).$$

The unbound variables of $\Phi(\mathbf{C}_{\varphi_1}, \dots, \mathbf{C}_{\varphi_n})$ are those of $\varphi_1, \dots, \varphi_n$.

3.8.5. *If Φ is a predicative formula and $\mathbf{C}_{\varphi_1}, \dots, \mathbf{C}_{\varphi_n}$ are collections without unbound variables then*

$$\mathfrak{N} \models \Phi([\varphi_1], \dots, [\varphi_n]) \leftrightarrow \mathfrak{M} \models \Phi(\mathbf{C}_{\varphi_1}, \dots, \mathbf{C}_{\varphi_n}).$$

◁ The proof proceeds by induction on the length of Φ on using the propositions of 3.8.3. ▷

3.8.6. The axioms of NCT other than predicative formulas have the shape

$$(Q_1 X)(Q_2 Y)(Q_3 Z)\Psi(X, Y, Z),$$

where $Q_1, Q_2 \in \{\forall, \forall^{\text{st}}\}$, $Q_3 \in \{\exists, \exists^{\text{st}}\}$, and Ψ is a predicative formula. We say that some formula of the above shape *holds in BST for classes* whenever to arbitrary formulas $\varphi_1(x, u_1, \dots, u_l)$ and $\varphi_2(x, v_1, \dots, v_m)$ of BST, assumed internal if the corresponding quantifiers are external, there is a formula $\varphi(x, w_1, \dots, w_n)$ of BST, internal if Q is an external quantifier, such that the formula

$$(Q_1 u_1) \dots (Q_1 u_l)(Q_2 v_1) \dots (Q_2 v_m)(Q_3 w_1) \dots (Q_3 w_n)\Psi(\mathbf{C}_{\varphi_1}, \mathbf{C}_{\varphi_2}, \mathbf{C}_{\varphi})$$

holds in BST. We also assume that the variables u_i, v_i , and w_i are absent in the construction of Ψ .

3.8.7. *Assume that Φ has the shape*

$$(Q_1 X)(Q_2 Y)(Q_3 Z)\Psi(X, Y, Z)$$

as in 3.8.6. *If Φ holds in BST for classes then Φ is valid in \mathfrak{N} .*

◁ Consider the case in which all quantifiers over classes in Φ are external. Take arbitrary \mathfrak{N} -standard elements $[\varphi_1], [\varphi_2] \in N$. Since Φ holds in BST for classes, it follows that there is an internal formula φ of the language L with \mathfrak{M} -standard parameters and a single unbound variable such that $\mathfrak{M} \models \Psi(\mathbf{C}_{\varphi_1}, \mathbf{C}_{\varphi_2}, \mathbf{C}_{\varphi})$. Hence, $\text{st}^N[\varphi]$ and $\mathfrak{N} \models \Psi([\varphi_1], [\varphi_2], [\varphi])$ by 3.8.5, which completes the proof. ▷

It is an easy matter to prove that the axioms of transfer, class formation, regularity, replacement, and idealization hold for classes in BST. The axioms of standardization, comprehension, and saturation deserve a special discussion.

We will use the definitions, notation, and above-proven facts about monads and ultrafilters which are preserved in BST. We also need the following theorem of [13].

3.8.8. Theorem. *To each formula Φ with two unbound variables there is an internal formula φ satisfying*

$$\begin{aligned} (\forall p)(\forall^{\text{st}}x)(\Phi(x, p) \leftrightarrow (\forall^{\text{st}}U \in \text{Ult})(p \in \nu(U) \rightarrow \varphi(x, U))) \\ \leftrightarrow (\exists^{\text{st}}U \in \text{Ult})(p \in \nu(U) \wedge \varphi(x, U)). \end{aligned}$$

3.8.9. Theorem. *The axiom of standardization of NCT holds in BST for classes.*

\triangleleft Let Φ be an arbitrary formula. Without loss of generality, we may assume that Φ has at most two unbound variables. Take some internal formula φ as suggested by Theorem 3.8.8. If p is a set and U is an ultrafilter such that $p \in \nu(U)$ then

$$(\forall^{\text{st}}x)(\Phi(x, p) \leftrightarrow \varphi(x, U)).$$

This proves the validity of the axiom of standardization in BST for classes since each set belongs to the nest of some ultrafilter. \triangleright

Let U be an ultrafilter. Assign

$$\begin{aligned} \text{dom}(U) &:= \{\text{dom}(u) : u \in U\}; \\ \text{im}(U) &:= \{\text{im}(u) : u \in U\}. \end{aligned}$$

It is easy to prove by transfer and idealization that $\text{dom}(U)$ and $\text{im}(U)$ are ultrafilters for every ultrafilter U . Moreover,

$$\begin{aligned} (a, b) \in \nu(U) &\rightarrow a \in \nu(\text{dom}(U)) \wedge b \in \nu(\text{im}(U)); \\ a \in \nu(\text{dom}(U)) &\rightarrow (\exists b \in \text{im}(U))((a, b) \in \nu(U)) \end{aligned}$$

for all sets a and b .

3.8.10. Theorem. *The axiom of comprehension holds in BST for classes.*

\triangleleft Let Φ be some formula with two unbound variables. By Theorem 3.8.8 there is an internal formula φ satisfying

$$\Phi(a, b) \leftrightarrow (\exists^{\text{st}}U \in \text{Ult})((a, b) \in \nu(U) \wedge \varphi(U)).$$

Put

$$\psi(V, W) := (\exists U \in \text{Ult})(\text{dom}(U) = V \wedge \text{im}(U) = W \wedge \varphi(U)).$$

Using the transfer principle of BST and Theorem 3.8.10, to each standard set A there is a standard set R such that

$$(\forall V \in \text{Ult}(A))((\exists W) \psi(V, W) \rightarrow (\exists W \in R) \psi(V, W)).$$

Put $Y := \bigcup \bigcup R$. By transfer and the properties of the nests of ultrafilters,

$$\begin{aligned} (\exists b)\Phi(a, b) &\rightarrow (\exists^{\text{st}}V \in \text{Ult}(A))(\exists^{\text{st}}W \in \text{Ult})(a \in \nu(V) \wedge \psi(V, W)) \\ &\rightarrow (\exists^{\text{st}}V \in \text{Ult}(A))(\exists^{\text{st}}W \in R)(a \in \nu(U) \wedge \psi(V, W)) \\ &\rightarrow (\exists b \in Y)(\exists^{\text{st}}U)((a, b) \in \nu(U) \wedge \varphi(U)) \\ &\rightarrow (\exists b \in Y)\Phi(a, b) \end{aligned}$$

for all $a \in A$.

Assume now that $\Psi(x, y, p)$ is an arbitrary formula. To complete the proof of the theorem it suffices to show that to each p and each standard X there is a set Y such that

$$(\forall x \in X)((\exists y)\Psi(x, y, p) \rightarrow (\exists y \in Y)\Psi(x, y, p)).$$

Distinguishing some standard sets X and p , put

$$\Phi(a, b) := (\exists x)(\exists p)(a = (x, p) \wedge \Psi(x, b, p)).$$

As we have proven already, there is a standard set Y such that the sought formula holds for all $p \in P$. We are done on applying the axiom of boundedness. \triangleright

3.8.11. Theorem. *The axiom of saturation holds in BST for classes.*

\triangleleft By Theorem 3.8.8 to each formula Φ with two unbound variables there is an internal formula φ satisfying

$$\Phi(x, p) \leftrightarrow (\exists^{\text{st}}U)((p, x) \in \nu(U) \wedge \varphi(U)).$$

Hence, given an arbitrary set p and using 3.7.9, we infer that

$$(\forall x)(\Phi(x, p) \rightarrow (\forall y \in \mu_p(x)) \Phi(y, p)),$$

so completing the proof. \triangleright

We have checked that all axioms of NCT hold in \mathfrak{N} , which ends the proof of Theorem 3.8.1.

3.8.12. COMMENTS.

(1) The fact that proper classes are not members of other classes clearly restricts the expressive means of NCT. In particular, it is impossible in NCT to formalize the construction of the nonstandard hull $E^\#$ of an internal normed space E to a full extend (cf. Chapter 6). Indeed, the members of $E^\#$ are the equivalence classes of the quotient of the class of limited members of E by the external relation of infinite proximity on E . Since these classes are both external, there is no class

that contains them as members. However, to view the nonstandard hull of E as the class comprising these equivalence classes we need furnish NCT with a stronger form of the axiom of choice which asserts for instance the possibility of ordering each semiset so that each subclass of this semiset has a least element. Such an order is a *strong well-ordering*. However, this cannot be implemented in NCT without a contradiction as shown in [14].

(2) We can prove that a class admits a strong well-ordering only if there is a bijection of this class to the semiset of standard elements of some standard set (the semiset of limited members of an internal normed space fails to possess this property).

(3) A class X has *standard size* provided that there are a function F and a class D satisfying $X = F^{\circ}D$. In this event, we may assume F internal and D standard. The article [14] contains the following

Theorem. *A semiset X admits a strong well-ordering if and only if X has standard size.*

(4) The theory NCT has tools allowing us to prove some proposition that is equivalent to the theorem of completeness of the nonstandard hull of an internal normed space. We mean the claim that each external S -fundamental sequence (e_n) of elements of E (i.e., a sequence indexed with standard naturals and such that $(\forall^{\text{st}}\varepsilon > 0)(\exists^{\text{st}}n_0)(\forall^{\text{st}}m, n > n_0)(\|e_n - e_m\| < \varepsilon)$) has S -limit in E (i.e., $(\exists e \in E)(\forall^{\text{st}}\varepsilon > 0)(\exists^{\text{st}}n_0)(\forall^{\text{st}}n > n_0)\|e_n - e\| < \varepsilon)$).

Analogously, NCT will allow us to formalize all considerations of Chapter 7 which deal with construction some topological groups as quotient groups of hyperfinite group by external normal subgroups.

3.9. Relative Internal Set Theory

In this section we consider the theory of relative internal sets within Nelson's internal set theory.

3.9.1. The presence of infinitesimals in analysis opens a way of constructing new concepts (and in fact of legitimizing the concepts that were refuted long ago) in order to study the classical objects of mathematical analysis. In particular, our new attractive acquisition is new mathematical concepts such as a microlimit of a finite sequence (cf. 2.3.4) or microcontinuity of a function at a point (cf. 4.4.5). These and similar concepts enable us to formulate the "infinitesimal" tests for limits (cf. 2.3.5), continuity (cf. 2.3.8 and 4.2.7), compactness (cf. 4.3.6), etc. which lie in the backgrounds of most applications.

However, all new tests presume the standard environment which is a limitation on their applicability. Moreover, even in case we use infinitesimal tests to studying

a standard object, we often face serious obstacles that stem from this limitation. We will proceed with two typical examples.

(1) Consider the infinitesimal test for the fact that

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} x_{nm} = a$$

for some standard double sequence $(x_{nm})_{n,m \in \mathbb{N}}$ and a standard real a . Apply the test of 2.3.5 to reformulate:

A standard number $a \in \mathbb{R}$ is the limit of a standard sequence (a_n) if and only if a is a microlimit of $a[N]$, with N an arbitrary unlimited natural number. In symbols,

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} x_{nm} = a \leftrightarrow (\forall N \approx +\infty) (\lim_{m \rightarrow +\infty} x_{N,m} \approx a).$$

However, the infinitesimal limit test is not applicable since the internal sequence $y_m := {}^*x_{N,m}$ fails to be standard in general. A similar obstacle appears in testing the double limit $\lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x, y)$ of a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ as well as in trying to find an infinitesimal presentation of the improper integral (cf. 2.3.6).

(2) Let us try now to supply an infinitesimal proof for the following familiar proposition:

The uniform limit of a sequence of bounded uniformly continuous function on a uniform space is a uniformly continuous function.

We thus consider a standard sequence $(f_n)_{n \in \mathbb{N}}$ of such functions converging uniformly to a standard function f on a set X . By the same reasons as in (1), if N is an unlimited natural and $x \in X$ then $f_N(x) \approx f(x)$, since $\sup\{|f_N(x) - f(x)| : x \in X\} \approx 0$. By 2.3.12 (cf. 4.4.6 (1)) it suffices to prove that $x' \approx x'' \rightarrow f(x') \approx f(x'')$ for all $x', x'' \in X$. Of course, $f(x') \approx f_N(x')$ and $f(x'') \approx f_N(x'')$. However, it is impossible to infer that $f_N(x') \approx f_N(x'')$ despite the fact that f is uniformly continuous since the infinitesimal uniform continuity test is not applicable to the (generally) nonstandard function f_N .

To obviate these obstacles we will introduce relatively standard elements. Informally speaking, a set relatively standard with respect to x is “more nonstandard” as compared with x . This reminds us of higher order infinitesimals.

3.9.2. We will consider the notion of relative standardness within IST.

Denote by $\text{Ffin}(f)$ the following: “ f is a function and each member of the image $\text{im}(f)$ of f is a finite set.” In symbols,

$$\text{Ffin}(f) := \text{Func}(f) \wedge (\forall x \in \text{im}(f)) \text{fin}(f(x)).$$

An element x is *feasible* (in symbols, $(\text{Su}(x))$) provided that $(\exists^{\text{st}} X)(x \in X)$. We introduced the predicate “ x is standard relative to y ” by the formula

$$x \text{ st } y := (\exists^{\text{st}} \varphi)(\text{Ffin}(\varphi) \wedge y \in \text{dom}(\varphi) \wedge x \in \varphi(y)).$$

It is worth recalling that $\text{fin}(x)$ means only that x is a finite set or, in other words, the cardinality or size of x is a member of ω ; i.e. x is a natural, possibly unlimited whenever x is a nonstandard set.

3.9.3. In the sequel we will use the following auxiliary proposition.

Let $\varphi(x, y)$ be a formula of ZFC. Assume that $(\forall x)(\exists! y)\varphi(x, y)$ is provable in ZFC. Then the formula $(\forall^{\text{st}}x)(\forall y)(\varphi(x, y) \rightarrow \text{St}(y))$ is provable in IST.

◁ This is immediate by transfer. ▷

In analogy with 3.3.4 we abbreviate some predicates as follows:

$$\begin{aligned} (\forall^{\text{st}}y x) \varphi &:= (\forall x) ((x \text{ is standard relative to } y) \rightarrow \varphi); \\ (\exists^{\text{st}}y x) \varphi &:= (\exists x) ((x \text{ is standard relative to } y) \wedge \varphi); \\ (\forall^{\text{st fin}}y x) \varphi &:= (\forall^{\text{st}}y x) (x \text{ is finite} \rightarrow \varphi); \\ (\exists^{\text{st fin}}y x) \varphi &:= (\exists^{\text{st}}y x) (x \text{ is finite} \wedge \varphi). \end{aligned}$$

3.9.4. The two-place predicate $x \text{ st } y$ possesses the following properties:

- (1) $x \text{ st } y \rightarrow \text{Su}(x) \wedge \text{Su}(y)$.
- (2) $x \text{ st } y \wedge y \text{ st } z \rightarrow x \text{ st } z$.
- (3) $x \text{ st } y \wedge \text{fin}(x) \rightarrow (\forall z \in x) z \text{ st } y$.
- (4) $\text{Su}(y) \wedge \text{St}(x) \rightarrow x \text{ st } y$.

In the last claim St is the familiar one-place predicate expressing “standardness” in IST; cf. 3.3.1.

◁ (1): Obvious.

(2): Assume that $x \in \varphi_1(y)$ and $y \in \varphi_2(z)$ where φ_1 and φ_2 are standard functions such that $\varphi_1(t)$ and $\varphi_2(t)$ are finite sets for all $t \in \text{dom}(\varphi_i)$. Let h stand for the function defined as

$$\text{dom}(h) := \text{dom}(\varphi_2), \quad h(t) := \{\varphi_1(u) : u \in \varphi_2(t) \cap \text{dom}(\varphi_1)\} \quad (t \in \text{dom}(\varphi_2)).$$

Clearly, $h(t)$ is a finite set for all $t \in \text{dom}(\varphi_2)$ since so is $\varphi_1(t)$ and $x \in h(z)$. By 3.9.3, h is a standard function, which finishes the proof of (2).

(3): Assume that ψ is a standard function such that $\psi(t)$ is finite for all $t \in \text{dom}(\psi)$. Take $x \in \psi(y)$. Appealing to 3.9.3 again, let g stand for the standard function defined by the rule

$$g(t) := \{v : v \in \psi(t) \wedge v \text{ is finite}\} \quad (t \in \text{dom}(g) := \text{dom}(\psi)).$$

It is clear that $z \in g(y)$, and we are done on observing that $g(t)$ is a finite set for all $t \in \text{dom}(\psi)$ as a union of finite sets.

(4): Let X be a standard set. This X exists since $\text{Su}(y)$. Define the function φ as $\varphi(t) := \{x\}$ ($t \in X$). This is a standard function since x is standard and $x \in \varphi(y)$. ▷

3.9.5. Relative Transfer Principle. *If A is an internal formula with the only unbound variables x, t_1, \dots, t_k ($k \geq 1$) then*

$$(\forall^{\text{st } \tau} t_1) \dots (\forall^{\text{st } \tau} t_k) ((\forall^{\text{st } \tau} x) A(x, t_1, \dots, t_k) \rightarrow (\forall x) A(x, t_1, \dots, t_k))$$

for an arbitrary feasible τ .

◁ For brevity, put $k := 1$. We have to prove validity of the following proposition in IST:

$$(\forall \tau)(\text{Su}(\tau) \rightarrow (\forall^{\text{st } \tau} t)((\forall^{\text{st } \tau} x) A(x, t) \rightarrow (\forall x) A(x, t))).$$

Since $\neg \text{Su}(t) \rightarrow \neg(t \text{ st } \tau)$, by 3.9.4 the formula in question amounts to the following

$$(\forall \tau)(\forall^{\text{st } \tau} t)((\forall^{\text{st } \tau} x) A(x, t) \rightarrow (\forall x) A(x, t)).$$

Translating this formula to the language of IST, infer

$$\begin{aligned} (\forall \tau)(\forall^{\text{st}} \varphi)(\forall t) (\tau \in \text{dom}(\varphi) \wedge t \in \varphi(\tau) \\ \rightarrow ((\forall^{\text{st}} \psi)(\forall x)(\tau \in \text{dom}(\psi) \wedge x \in \psi(\tau) \rightarrow A(x, t)) \\ \rightarrow (\forall y) A(y, t))). \end{aligned}$$

Here and in what follows, we let φ and ψ stand for some functions whose values are finite sets while Φ is some set of these functions. Proceed by applying the Nelson algorithm 3.3.15 so arriving at the new equivalent record (in predicate calculus):

$$\begin{aligned} (\forall^{\text{st}} \varphi) (\forall \tau, t) (\exists^{\text{st}} \psi)(\tau \in \text{dom}(\varphi) \wedge t \in \varphi(\tau) \\ \wedge (\forall x)(\tau \in \text{dom}(\psi) \wedge x \in \psi(\tau) \rightarrow A(x, t)) \rightarrow (\forall y) A(y, t)). \end{aligned}$$

By idealization, the last formula amounts to the following

$$\begin{aligned} (\forall^{\text{st}} \varphi)(\exists^{\text{st fin}} \Phi)(\forall \tau) (\forall t)(\tau \in \text{dom}(\varphi) \wedge t \in \varphi(\tau) \\ \wedge (\forall \psi \in \Phi)(\forall x)(\tau \in \text{dom}(\psi) \wedge x \in \psi(\tau) \rightarrow A(x, t)) \rightarrow (\forall y) A(y, t)). \end{aligned}$$

By transfer, we omit the superscript st of the first two quantifiers and reduce the task to validating the following formula in ZFC:

$$\begin{aligned} (\forall \varphi)(\exists^{\text{fin}} \Phi)(\forall \tau, t)(\tau \in \text{dom}(\varphi) \wedge t \in \varphi(\tau) \\ \wedge (\forall \psi \in \Phi)(\forall x)(\tau \in \text{dom}(\psi) \wedge x \in \psi(\tau) \rightarrow A(x, t)) \\ \rightarrow (\forall y) A(y, t)). \end{aligned}$$

Take some function φ , and arrange the singleton $\Phi := \{\psi\}$ where ψ is defined as follows: Put $M := \bigcup \text{im}(\varphi)$ and $M_1 := \{t \in M : (\exists y) \neg A(y, t)\}$. The axioms of

ZFC implies that there is a function h such that $\text{dom}(h) = M_1$ and $\neg A(h(t), t)$ for all $t \in \text{dom}(h)$. Assign $\text{dom}(\psi) := \text{dom}(\varphi)$ and $\psi(\alpha) := \{h(\nu) : \nu \in \varphi(\alpha) \cap M_1\}$ ($\alpha \in \text{dom}(\varphi)$). If $\varphi(\alpha) \cap M_1 = \emptyset$ then $\psi(\alpha) := \emptyset$. Note that $\psi(\alpha)$ is finite since so is $\varphi(\alpha)$. Distinguish $\tau \in \text{dom}(\varphi)$ and $t \in \varphi(\tau)$. To validate the formula of ZFC under study, it suffices now to prove the following implication:

$$(\forall x \in \psi(\tau)) A(x, t) \rightarrow (\forall y) A(y, t).$$

If $t \in M - M_1$ then this implication is valid, whereas if $t \in M_1$, the premise of this implication is false. Indeed, if $x = h(t)$ then $x \in \psi(\tau)$ since $t \in \varphi(\tau) \cap M_1$, whereas $A(h(t), t)$ is false. \triangleright

3.9.6. Relative Idealization Principle. *Let $B(x, y)$ be an internal formula possibly possessing unbound variables other than x and y . Then*

$$(\forall^{\text{st } \tau \text{ fin}} z)(\exists x)(\forall y \in z) B(x, y) \leftrightarrow (\exists x)(\forall^{\text{st } \tau} y) B(x, y)$$

for every feasible τ .

\triangleleft As in 3.9.5, we let φ and ψ denote some functions whose values are finite sets, while Φ and Ψ are some sets of these functions. Note first that to prove the claim it suffices to demonstrate the implication \rightarrow since the reverse implication follows from 3.9.4 (3).

Assume for definiteness that B contains one more unbound variable t . Then the task consists in validating the formula

$$\begin{aligned} (\forall \tau)(\text{Su}(\tau) \rightarrow (\forall t)(\forall^{\text{st } \tau \text{ fin}} x)(\exists y)(\forall z \in x) B(z, y, t) \\ \rightarrow (\exists u)(\forall^{\text{st } \tau} v) B(v, u, t))). \end{aligned}$$

Appealing to 3.9.4 again, observe that the above formula amounts to the following

$$(\forall \tau)(\forall t)((\forall^{\text{st } \tau \text{ fin}} x)(\exists y)(\forall z \in x) B(z, y, t) \rightarrow (\exists u)(\forall^{\text{st } \tau} v) B(v, u, t)).$$

Rephrasing this in the language of IST, replace the predicate $x \text{ st } \tau$ with the equivalent fragment in the language of IST to obtain

$$\begin{aligned} (\forall \tau)(\forall t)((\forall^{\text{st}} \varphi)(\forall^{\text{fin}} x)(\tau \in \text{dom}(\varphi) \wedge x \in \varphi(\tau) \\ \rightarrow (\exists y)(\forall z \in x) B(z, y, t)) \\ \rightarrow (\exists u)(\forall^{\text{st}} \Psi)(\forall v)(\tau \in \text{dom}(\psi) \wedge v \in \psi(\tau) \rightarrow B(v, u, t))) \\ \leftrightarrow \\ (\forall \tau)(\forall t)((\forall^{\text{st}} \varphi)(\forall^{\text{fin}} x)(\tau \in \text{dom}(\varphi) \wedge x \in \varphi(\tau) \end{aligned}$$

$$\begin{aligned}
& \rightarrow (\exists y)(\forall z \in x) B(z, y, t)) \\
& \rightarrow (\forall^{\text{st}} \text{fin} \Psi)(\exists u)(\forall \psi \in \Psi)(\forall v)(\tau \in \text{dom}(\psi) \wedge v \in \psi(\tau) \\
& \rightarrow B(v, u, t))) \\
& \leftrightarrow \\
& (\forall^{\text{st}} \text{fin} \Psi)(\exists^{\text{st}} \text{fin} \Phi)(\forall \tau)(\forall t)((\forall \varphi \in \Phi)(\forall^{\text{fin}} x) \\
& (\tau \in \text{dom}(\varphi) \wedge x \in \varphi(\tau) \\
& \rightarrow (\exists y)(\forall z \in x) B(z, y, t)) \rightarrow (\exists u)(\forall \psi \in \Psi)(\forall v) \\
& (\tau \in \text{dom}(\psi) \wedge v \in \psi(\tau) \rightarrow B(v, u, t))).
\end{aligned}$$

In much the same way as in the proof of 3.9.5, we may omit the superscript *st* in the first two quantifiers by transfer, reducing the matter again to validating the so-obtained formula of ZFC. Take an arbitrary finite set of functions Ψ and define Φ by the rule $\Phi := \{\varphi\}$ where

$$\text{dom}(\varphi) := \bigcup_{\psi \in \Psi} \text{dom}(\psi), \quad \varphi(\alpha) := \left\{ \bigcup \psi(\alpha) : \psi \in \Psi, \alpha \in \text{dom}(\psi) \right\}.$$

Note that $\varphi(\alpha)$ is a finite set. Distinguish some arbitrary τ and t . If $\tau \notin \text{dom}(\varphi)$ then $\tau \notin \text{dom}(\psi)$ for all $\psi \in \Psi$, implying that the formula in question is valid. In case $\tau \in \text{dom}(\varphi)$, we are left with validating the implication

$$\begin{aligned}
& (\exists y) \left(\forall z \in \bigcup \{\psi(\tau) : \psi \in \Psi, \tau \in \text{dom}(\psi)\} \right) B(z, y, t) \\
& \rightarrow (\exists u)(\forall \psi \in \Psi)(\forall v)(\tau \in \text{dom}(\psi) \wedge v \in \psi(\tau) \rightarrow B(v, u, t)).
\end{aligned}$$

Rephrase the premise of this implication as

$$\begin{aligned}
& (\exists y)(\forall z \in \bigcup \{\psi(\tau) : \psi \in \Psi, \tau \in \text{dom}(\psi)\}) B(z, y, t) \\
& \leftrightarrow \\
& (\exists y)(\forall z)(\exists \psi \in \Psi)(\tau \in \text{dom}(\psi) \wedge z \in \psi(\tau) \rightarrow B(z, y, t)) \\
& \leftrightarrow \\
& (\exists y)(\forall z)(\forall \psi \in \Psi)(\tau \in \text{dom}(\psi) \wedge z \in \psi(\tau) \rightarrow B(z, y, t)).
\end{aligned}$$

It is now clear that the premise of this implication is equivalent to its conclusion. \triangleright

3.9.7. We now list a few simple corollaries to 3.9.5 and 3.9.6.

- (1) *In the context of 3.9.3, let x be some τ -standard element, and let y satisfy $A(x, y)$. Then y is a τ -standard element too.*

◁ Immediate by relative transfer. ▷

(2) For a τ -standard set x to be finite it is necessary and sufficient that x consist of τ -standard elements.

◁ The implication \rightarrow coincides with 3.9.4 (3). To prove the reverse implication, rewrite the claim as $(\forall u \in x)(\exists^{\text{st } \tau} v)(u = v)$. By relative idealization, the last formula may be rephrased as $(\exists^{\text{st } \tau} \text{fin } V)(\forall u \in x)(\exists v \in V)(u = v)$ which means that $x \subset V$ and x is finite since so is V . ▷

(3) $\text{fin}(x) \rightarrow |x| \text{st } \tau$.

◁ Immediate from (1). ▷

3.9.8. We now show that it is impossible to prove any analog of 3.9.5 and 3.9.6 for standardization. To this end, we rewrite the alias “relative standardization principle” as follows:

(1) $(\forall^{\text{st } \tau} x)(\exists^{\text{st } \tau} y)(\forall^{\text{st } \tau} z)(z \in y \leftrightarrow z \in x \wedge C(z))$,

where $C(z)$ is a formula of IST which may possess unbound variable other than z . We will demonstrate that 3.9.8 (1) leads to a contradiction even if $C(z)$ satisfies the following extra requirement: Each occurrence of the predicate st in $C(z)$ has the shape $\cdot \text{st } \tau$, whereas the unary predicate $\text{st}(x)$ is absent from $C(z)$.

Indeed, the existence of the standard part ${}^{\circ}t$ of an arbitrary limited real $t \in \mathbb{R}$ follows in IST by standardization, cf. 2.2.16. The same arguments may be repeated for the τ -standard part operation on assuming that 3.9.8 (1) holds. We now fill in details.

Let τ be an arbitrary feasible internal set. A real $x \in \mathbb{R}$ is τ -infinitesimal (in symbols, $x \overset{\tau}{\approx} 0$) provided that $|x| \leq \varepsilon$ for every τ -standard strictly positive $\varepsilon \in \mathbb{R}$. From 3.9.8 (1) with the appropriate $C(z)$ satisfying the above condition it ensues that

$$(\forall \tau)(\forall t \in \mathbb{R})((\exists^{\text{st } \tau} u \in \mathbb{R})(|t| < u) \rightarrow (\exists^{\text{st } \tau} v \in \mathbb{R})(|t - v| \overset{\tau}{\approx} 0)).$$

To see the falsity of 3.9.8 (1) it suffices to appeal to the following proposition.

3.9.9. There are an infinitely large natural N and $x \in [0, 1]$ such that y fails to be N -standard whenever y is N -infinitely close to x .

◁ The proof is given below in 4.6.15. ▷

3.9.10. Closing the current section, we briefly present the axiomatic theory RIST of relative internal sets. The language of this theory results from the language of Zermelo–Fraenkel set theory by supplementing it with a sole two-place predicate st . As before, we read the expression $x \text{st } y$ as “ x is standard relative to y .” A formula of RIST is internal if it contains no occurrences of the predicate st . In much the same way as in 3.9.3 we define the external quantifiers $\forall^{\text{st } \alpha}$, $\exists^{\text{st } \alpha}$, $\forall^{\text{st fin } \alpha}$,

$\exists^{\text{st fin}} \alpha$. The axioms of RIST contain all axioms of ZFC. Moreover, the predicate st obeys the following three axioms:

- (1) $(\forall x) x \text{ st } x$;
- (2) $(\forall x)(\forall y) x \text{ st } y \vee y \text{ st } x$;
- (3) $(\forall x)(\forall y)(\forall z)(x \text{ st } y \wedge y \text{ st } z \rightarrow x \text{ st } z)$.

Moreover, RIST, in analogy with IST, includes three new axiom-schemata. The axiom-schemata of transfer and idealization are the same as in 3.9.4 and 3.9.5, whereas we must restrict the class of formulas in the axiom-schema of standardization in accord with 3.9.6.

3.9.11. Axiom-Schema of Transfer. *If $\varphi(x, t_1, \dots, t_k)$ is an internal formula with free variables x, t_1, \dots, t_k and τ a fixed set then*

$$(\forall^{\text{st } \tau} t_1) \dots (\forall^{\text{st } \tau} t_k) ((\forall^{\text{st } \tau} x) \varphi(x, t_1, \dots, t_k) \rightarrow (\forall x) \varphi(x, t_1, \dots, t_k)).$$

3.9.12. Axiom-Schemata of Idealization. *Let $\varphi(x_1, \dots, x_k, y)$ be an internal formula with unbound variables x_1, \dots, x_k, y and possibly other unbound variables. Assume that τ_1, \dots, τ_k are given and β is not standard relative to (τ_1, \dots, τ_k) . Then*

- (1) **Restricted Idealization Principle:**
 $(\forall^{\text{st } \tau_1 \text{ fin}} z_1) \dots (\forall^{\text{st } \tau_k \text{ fin}} z_k)$
 $(\exists^{\text{st } \beta} y)(\forall x_1 \in z_1) \dots (\forall x_k \in z_k) \varphi(x_1, \dots, x_k, y)$
 $\leftrightarrow (\exists^{\text{st } \beta} y)(\forall^{\text{st } \tau_1} x_1) \dots (\forall^{\text{st } \tau_k} x_k) \varphi(x_1, \dots, x_k, y)$.
- (2) **Unrestricted Idealization Principle:**
 $(\forall^{\text{st } \tau \text{ fin}} z_1) \dots (\forall^{\text{st } \tau \text{ fin}} z_k)$
 $(\exists y)(\forall x_1 \in z_1) \dots (\forall x_k \in z_k) \varphi(x_1, \dots, x_k, y)$
 $\leftrightarrow (\exists y)(\forall^{\text{st } \tau_1} x_1) \dots (\forall^{\text{st } \tau_k} x_k) \varphi(x_1, \dots, x_k, y)$.

3.9.13. To formulate the axiom-schema of standardization, we introduce the class \mathcal{F}_τ of τ -external formulas, with τ a distinguished set. If \mathcal{F} is a class of formulas of RIST then \mathcal{F}_τ is defined as the least subclass of \mathcal{F} meeting the conditions:

- (1) Each atomic formula $x \in y$, with x and y variables or constants, belongs to \mathcal{F}_τ ;
- (2) If some formulas φ and ψ belong to \mathcal{F}_τ then the formulas $\neg\varphi$ and $\varphi \rightarrow \psi$ belong to \mathcal{F}_τ too;
- (3) If a formula $\varphi(x, y)$ belongs to \mathcal{F}_τ then the formula $(\exists y) \varphi(x, y)$ belongs to \mathcal{F}_τ as well;
- (4) If a formula $\varphi(x, y)$ belongs to \mathcal{F}_τ and β is a set such that the set τ is standard relative to β , then the formula $(\exists^{\text{st } \beta} y) \varphi(x, y)$ belongs to \mathcal{F}_τ .

3.9.14. Axiom-Schema of Standardization. If τ is a given set and φ is some τ -external formula then

$$(\forall^{\text{st } \tau} y)(\exists^{\text{st } \tau} z)(\forall^{\text{st } \tau} t)(t \in z \leftrightarrow (t \in y \wedge \varphi(t))).$$

3.9.15. Theorem. RIST is conservative over ZFC.

3.9.16. COMMENTS.

(1) The content of 3.9.2–3.9.9 is taken from Gordon’s article [504]; also see [146]. Since the relative standardization principle is not valid, we may conclude that the standardization principle of IST is not a consequence of the other axioms of this theory (for details, see [504, 146]).

(2) The definitions and basic properties of relative standard elements are presented within IST. However, all these results are naturally valid in an arbitrary nonstandard universe maintaining the idealization principle of IST. Applying these tools, we must surely bear in mind the particularity of the classical stance of infinitesimal analysis, cf. 3.5.2–3.5.12.

(3) The axiomatic theory RIST, as presented in 3.9.10–3.9.14, was propounded by Péraire [395]. The last article contains Theorem 3.9.15. Prior to this, Péraire carried out an extension of IST (consistent with ZFC) by appending a sequence of the undefinable predicates $\text{St}_p(x)$ (read: x is “standard to the power of” $1/p$); cf. [393]. The articles [394] and [396] contain other results in this direction.

(4) The articles [141, 146] also consider simpler version of the concept of relative standardness. Namely, we introduce the relation x is *strongly standard relative to* τ or x is *τ -strongly-standard* by the formula

$$x \text{ sst } \tau := (\exists^{\text{st}} \varphi) (\text{Fnc } (\varphi) \wedge \tau \in \text{dom}(\varphi) \wedge x = \varphi(\tau)).$$

Clearly, $x \text{ sst } y \rightarrow x \text{ st } y$. However, the converse fails in general, see. [146].

(5) The relative transfer principle, as well as 3.9.4 (1, 2, 4), remains valid provided that we replace all occurrences of the predicate $\cdot \text{st} \cdot$ with $\cdot \text{sst} \cdot$. The relative idealization principle reduces to the implication \rightarrow .

(6) The following hold: (cf. [146]):

$$(\exists N \in \omega) (\exists n < N) (\neg n \text{ sst } N).$$

◁ It suffices to check the falsity in IST of the following

$$(\forall N \in \omega) (\forall n < N) (\exists^{\text{st}} \varphi \in \omega^\omega) (n = \varphi(N)).$$

Rephrasing this formula by idealization and transfer, we obtain

$$(\exists \Phi \in \mathcal{P}_{\text{fin}}(\omega))(\forall n, N \in \omega)(n < N \rightarrow (\exists \varphi \in \Phi)(n = \varphi(N))).$$

We now prove that the negation of the last formula is true in ZFC.

To this end, let $\Phi = \{\varphi_1, \dots, \varphi_k\}$ be some finite set of functions $\varphi_l : \omega \rightarrow \omega$. Take $N > k$. Clearly, there is some n satisfying $n \in \{0, 1, \dots, N-1\} - \{\varphi_1(N), \dots, \varphi_k(N)\}$. These n and N satisfy the formula $n < N \wedge (\forall \varphi \in \Phi)(n \neq \varphi(N))$, which ends the proof. \triangleright

Chapter 4

Monads in General Topology

The set-theoretic stance of mathematics has provided us with the environment known today as general topology for studying continuity and proximity since the beginning of the 20th century.

Considering the microstructure of the real axis, we have already seen that the collection of infinitesimals arises within infinitesimal analysis as a monad, i.e., the external intersection of all standard elements of the neighborhood filter of zero in the only separated topology agreeable with the algebraic structure of the field of reals.

We may say that the notion of the monad of a filter synthesizes to some extent the topological idea of proximity and the analytical idea of infinitesimality. Interplay between these ideas is the main topic of the current chapter.

We focus attention on the most elaborate ways of studying classical topological concepts and constructions that surround compactness and rest on the idealization principle we accept in nonstandard set theory.

The contribution of the new approach to the topic we discuss resides basically in evoking the crucial notion of a nearstandard point. The corresponding test for a standard space to be compact, consisting in nearstandardness of every point, demonstrates the meaning and significance of the concept of nearstandardness which translates the conventional notion of compactness from whole spaces to individual points. This technique of individualization is a powerful and serviceable weapon in the toolbox of infinitesimal analysis.

It is worth observing that we conform mainly to the neoclassical credo of infinitesimal analysis in this chapter, and so exposition proceeds in the standard environment unless otherwise stated.

4.1. Monads and Filters

The simplest example of a filter is well-known to be the collection of supersets of

a nonempty set. Infinitesimal analysis enables us to approach an arbitrary standard filter in much the same manner viewing this filter as standardization of the collection of supersets of an appropriate external set, the monad of the filter. The method for introducing these monads and studying their simplest properties constitutes the topic of this section.

4.1.1. Let X be a standard set and let \mathcal{B} be a standard filterbase on X . In particular, $\mathcal{B} \neq \emptyset$, $\mathcal{B} \subset \mathcal{P}(X)$, $\emptyset \notin \mathcal{B}$, and $B_1, B_2 \in \mathcal{B} \rightarrow (\exists B \in \mathcal{B})(B \subset B_1 \cap B_2)$. The symbol $\mu(\mathcal{B})$ denotes the *monad* of \mathcal{B} , i.e., the external set defined as follows:

$$\mu(\mathcal{B}) := \bigcap \{B : B \in {}^\circ\mathcal{B}\}.$$

4.1.2. An internal set A is a superset of some standard element of a standard filterbase \mathcal{B} if and only if A includes the monad $\mu(\mathcal{B})$ of \mathcal{B} .

◁ If $A \supset B$ and $B \in {}^\circ\mathcal{B}$ then $A \supset \mu(\mathcal{B})$ by definition. Conversely, if $A \supset \mu(\mathcal{B})$ then by idealization there is an internal set $B \in \mathcal{B}$ such that $B \subset \mu(\mathcal{B})$, and so $A \supset B$. ▷

4.1.3. Each standard filter \mathcal{F} is the standardization of the principal external filter of supersets of the monad $\mu(\mathcal{F})$.

◁ In symbols, the claim reads:

$$(\forall^{\text{st}} A)((A \in \mathcal{F}) \leftrightarrow (A \supset \mu(\mathcal{F}))).$$

The last equivalence is obviously a consequence of 4.1.2. ▷

4.1.4. The monad $\mu(\mathcal{F})$ of a standard filter \mathcal{F} is an internal set if and only if $\mu(\mathcal{F})$ is a standard set. In this event \mathcal{F} is the standard filter of supersets of $\mu(\mathcal{F})$.

◁ If $\mu(\mathcal{F})$ is an internal set then, using 4.1.3 and the idealization principle, we find

$$\begin{aligned} & (\exists A)(\forall^{\text{st}} F)(F \in \mathcal{F}) \\ \leftrightarrow & (F \supset A) \leftrightarrow (\forall^{\text{st fin}} \mathcal{U})(\exists A)(\forall F \in \mathcal{U})(F \in \mathcal{F} \leftrightarrow F \supset A) \\ \leftrightarrow & (\forall^{\text{st}} U)(\exists A)(U \in \mathcal{F} \leftrightarrow U \supset A). \end{aligned}$$

By transfer, we conclude that \mathcal{F} is the filter of supersets of some set A . Since such a set A is unique; therefore, $A = \mu(\mathcal{F})$ and, moreover, A is a standard set. ▷

4.1.5. Given a standard filterbase \mathcal{B} , we call the members of $\mu(\mathcal{B})$ *infinitesimal* or *distant*, or *remote*, or *astray* (relative to \mathcal{B}). Analogously, an element B in \mathcal{B} such that $B \subset \mu(\mathcal{B})$ is also called an infinitesimal or distant, or remote, or astray member of \mathcal{B} . The collection of infinitesimal members of \mathcal{B} is denoted by ${}^a\mathcal{B}$.

4.1.6. EXAMPLES.

(1) The monad $\mu(\mathbb{R})$ is the monad of the neighborhood filter of zero in the natural topology on \mathbb{R} .

(2) Let \mathcal{B} be a filterbase and let $\text{fil } \mathcal{B}$ be the filter with filterbase \mathcal{B} , i.e., \mathcal{B} is the collection of supersets of members of \mathcal{B} . In symbols,

$$\text{fil } \mathcal{B} := \{F \subset X : (\exists B \in \mathcal{B})(B \subset F)\}.$$

By transfer, if \mathcal{B} is a standard filterbase (on a standard set X) then $\text{fil } \mathcal{B}$ is also a standard filter. In this case $\mu(\mathcal{B}) = \mu(\text{fil } \mathcal{B})$.

Observe that in the sequel we will conveniently deal with the monad of an arbitrary internal filter \mathcal{F} which is determined in a perfect analogy with the above: $\mu(\mathcal{F}) := \bigcap \circ \mathcal{F}$. Observe also that the monad of a filter \mathcal{F} on a standard set is always an external superset of some internal element of \mathcal{F} .

(3) Let Ξ be a standard direction, i.e., a nonempty directed set. By idealization, Ξ contains an internal element that dominates all standard points of Ξ . Such a member of Ξ is called *distant*, or *remote*, *infinitely large*, or *unlimited*, or *astray* in Ξ .

Consider a standard base for the tail filter $\mathcal{B} := \{[\xi, \rightarrow) := \{\eta \in \Xi : \eta \geq \xi\} : \xi \in \Xi\}$ of Ξ . By definition, $\eta \in \mu(\mathcal{B}) \leftrightarrow (\forall^{\text{st}} \xi \in \Xi) \eta \geq \xi$, i.e., the monad of the tail filter of Ξ naturally comprises remote elements of Ξ . We will use the notation ${}^a\Xi := \mu(\mathcal{B})$.

(4) Let \mathcal{E} be a standard cover of a standard set X , i.e., $X \subset \bigcup \mathcal{E}$. Consider the family $\Xi(\mathcal{E})$ of standard finite unions of elements of \mathcal{E} . Clearly, $\Xi(\mathcal{E}) := * \{\bigcup \mathcal{E}_0 : \mathcal{E}_0 \in \mathcal{P}_{\text{st fin}}(\mathcal{E})\}$, where $\mathcal{P}_{\text{st fin}}(\mathcal{E})$ is the set of standard finite subsets of \mathcal{E} . The external collection of remote elements of $\Xi(\mathcal{E})$ is called the *monad* of \mathcal{E} and is denoted by $\mu(\mathcal{E})$. Hence,

$$\mu(\mathcal{E}) = \bigcup \{E : E \in \circ \mathcal{E}\}.$$

By analogy, we define the monad of each upward-filtered family.

(5) Assume given a correspondence $f \subset X \times Y$ and a filterbase \mathcal{F} on X such that f meets \mathcal{F} , i.e., $(\forall F \in \mathcal{F}) \text{dom}(f) \cap F \neq \emptyset$. As usual, we put

$$f(\mathcal{F}) := \{B \subset Y : (\exists F \in \mathcal{F})(B \supset f(F))\}.$$

Therefore, $f(\mathcal{F})$ is a filter on Y , called the *image* of \mathcal{F} under the correspondence f . Assuming the standard environment, i.e., on supposing that X , Y , f , and \mathcal{F} are

standard objects, we recall the idealization principle to obtain

$$\begin{aligned}
y \in \mu(f(\mathcal{F})) &\leftrightarrow (\forall^{\text{st}} B \in f(\mathcal{F}))(y \in B) \leftrightarrow (\forall^{\text{st}} F \in \mathcal{F})(y \in f(F)) \\
&\leftrightarrow (\forall^{\text{st}} F \in \mathcal{F})(\exists x)(x \in F \wedge y \in f(x)) \\
&\leftrightarrow (\forall^{\text{st fin}} \mathcal{F}_0 \subset \mathcal{F})(\exists x)(\forall F \in \mathcal{F}_0)(x \in F \wedge y \in f(x)) \\
&\leftrightarrow (\exists x)(\forall^{\text{st}} F \in \mathcal{F})(x \in F \wedge y \in f(x)) \\
&\leftrightarrow (\exists x \in \mu(\mathcal{F}))(y \in f(x) \leftrightarrow y \in f(\mu(\mathcal{F}))).
\end{aligned}$$

Therefore, the image of the monad of a filter is the monad of the image of this filter:

$$\mu(f(\mathcal{F})) = f(\mu(\mathcal{F})).$$

We now assume that \mathcal{G} is a filterbase on Y such that f^{-1} meets \mathcal{G} . Consider the *preimage* or *inverse image* $f^{-1}(\mathcal{G})$ of \mathcal{G} under f (i.e., the image of this filter under the correspondence f^{-1}). Obviously, $\mu(f^{-1}(\mathcal{G})) = f^{-1}(\mu(\mathcal{G}))$.

It is worth noting that the last relation may be proved without “saturation.” Indeed, recalling definitions, we simply deduce

$$\mu(f^{-1}(\mathcal{G})) = \bigcap_{G \in {}^\circ\mathcal{G}} f^{-1}(G) = f^{-1}\left(\bigcap_{G \in {}^\circ\mathcal{G}} G\right) = f^{-1}(\mu(\mathcal{G})),$$

i.e., the monad of the inverse image of a filter is the inverse image of the monad of the initial filter. Observe that, proving this, we use the possibility of dealing with the inverse image under f of an arbitrary external subset of Y .

4.1.7. Let \mathcal{B}_1 and \mathcal{B}_2 be two standard filterbases on some standard set. Then

$$\text{fil } \mathcal{B}_1 \supset \text{fil } \mathcal{B}_2 \leftrightarrow \mu(\mathcal{B}_1) \subset \mu(\mathcal{B}_2).$$

$\triangleleft \rightarrow$: If B_2 is standard and $B_2 \supset \mu(\mathcal{B}_2)$ then, by 4.1.2, $B_2 \in \text{fil } \mathcal{B}_2$ and so $B_2 \in \text{fil } \mathcal{B}_1$. Therefore, $B_2 \supset \mu(\mathcal{B}_1)$, yielding $\mu(\mathcal{B}_1) \subset \mu(\mathcal{B}_2)$.

\leftarrow : Let F_2 be a standard element of $\text{fil } \mathcal{B}_2$, i.e., a superset of some standard $B_2 \in \mathcal{B}_2$. By hypothesis, B_2 includes the monad $\mu(\mathcal{B}_1)$. By 4.1.2, $B_2 \in \text{fil } \mathcal{B}_1$ and so $F_2 \in \text{fil } \mathcal{B}_1$. It suffices to appeal to the transfer principle. \triangleright

4.1.8. Consider a mapping $f : X \rightarrow Y$. Suppose that \mathcal{A} is a filterbase on X and \mathcal{B} is a filterbase on Y . Assuming the standard environment, the following are valid:

- (1) $f(\mathcal{A}) \supset \text{fil } \mathcal{B}$;
- (2) $f^{-1}(\mathcal{B}) \subset \text{fil } \mathcal{A}$;
- (3) $\mu(f(\mathcal{A})) \subset \mu(\mathcal{B})$;
- (4) $f(\mu(\mathcal{A})) \subset \mu(\mathcal{B})$.

◁ To prove the equivalence (1) \leftrightarrow (2), proceed as follows:

$$\begin{aligned} f(\mathcal{A}) \supset \text{fil } \mathcal{B} &\leftrightarrow (\forall B \in \mathcal{B})(\exists A \in \mathcal{A})(f(A) \subset B) \\ &\leftrightarrow (\forall B \in \mathcal{B})(\exists A \in \mathcal{A}) A \subset f^{-1}(B) \leftrightarrow f^{-1}(\mathcal{B}) \subset \text{fil } \mathcal{A}. \end{aligned}$$

Equivalence of (1) and (3) ensues from 4.1.7. We now recall 4.1.6 (5) to infer

$$\begin{aligned} f(\mu(\mathcal{A})) \subset \mu(\mathcal{B}) &\leftrightarrow \mu(\mathcal{A}) \subset f^{-1}(\mu(\mathcal{B})) \\ &\leftrightarrow \mu(\mathcal{A}) \subset \mu(f^{-1}(\mathcal{B})) \leftrightarrow f^{-1}(\mathcal{B}) \subset \text{fil } \mathcal{A}, \end{aligned}$$

which completes the proof. ▷

4.1.9. Using the classical stance, we may simplify the statement of 4.1.8. This is done by omitting the words “assuming the standard environment” while restating 4.1.8 (4) as $*f(\mu(\mathcal{A})) \subset \mu(\mathcal{B})$, with $*$ denoting the robinsonian standardization. It is in common parlance to presume $f := *f$, which leads to the most comprehensible and easily-memorized formulation. The same formulation is also often within the neoclassical and radical stances. In other words, if infinitesimal analysis is used as a technique for exploring the von Neumann universe, we assume the explicit parameters standard unless otherwise stated, abbreviating the term “internal set” to “set.” This convenient agreement correlates obviously with the qualitative views of standard objects.

In the sequel, we will continue to pursue a freelance attitude, omitting any indication to the type of sets under study whenever this entails no confusion.

4.1.10. *The following hold:*

- (1) *The filters \mathcal{F}_1 and \mathcal{F}_2 have the least upper bound if and only if $\mu(\mathcal{F}_1) \cap \mu(\mathcal{F}_2) \neq \emptyset$;*
- (2) *Given an upper bounded set of filters \mathcal{E} , we have*

$$\mu(\text{sup } \mathcal{E}) = \bigcap \{\mu(\mathcal{F}) : \mathcal{F} \in {}^\circ\mathcal{E}\},$$

i.e., the monad of the intersection of filters is the intersection of their monads.

◁ (1): This is immediate from 4.1.7.

(2): Note first that for $\mathcal{F} \in {}^\circ\mathcal{E}$ we have $\mathcal{F} \leq \text{sup } \mathcal{E}$ and so $\mu(\text{sup } \mathcal{E}) \subset \mu(\mathcal{F})$. This implies the inclusion $\mu(\text{sup } \mathcal{E}) \subset \bigcap \{\mu(\mathcal{F}) : \mathcal{F} \in {}^\circ\mathcal{E}\}$. Assume now that $F \in {}^\circ\text{sup } \mathcal{E}$. By the definition of filter, there is a standard finite set $\mathcal{E}_0 \subset \mathcal{E}$ such that $F \in \text{sup } \mathcal{E}_0$. Using 4.1.3 and (1), we deduce $F \supset \mu(\text{sup } \mathcal{E}_0) = \bigcap \{\mu(\mathcal{F}) : \mathcal{F} \in \mathcal{E}_0\}$. Finally,

$$\mu(\text{sup } \mathcal{E}) \supset \bigcap \{\mu(\mathcal{F}) : \mathcal{F} \in \mathcal{E}_0, \mathcal{E}_0 \in \mathcal{P}_{\text{st fin}}(\mathcal{E})\} = \bigcap \{\mu(\mathcal{F}) : \mathcal{F} \in {}^\circ\mathcal{E}\},$$

which completes the proof. ▷

4.1.11. Assume that \mathcal{A} is an ultrafilter on X , i.e., an inclusion maximal member of the set $\mathcal{F}(X)$ of filters on X . Assume further that \mathcal{F} is another filter on X , i.e., $\mathcal{F} \in \mathcal{F}(X)$. Then either $\mu(\mathcal{A}) \cap \mu(\mathcal{F}) = \emptyset$ or $\mu(\mathcal{A}) \subset \mu(\mathcal{F})$.

◁ If $\mu(\mathcal{A}) \cap \mu(\mathcal{F}) \neq \emptyset$, then, by 4.1.10 (1), we may find the least upper bound $\mathcal{A} \vee \mathcal{F} = \mathcal{A}$. Hence, $\mathcal{F} \subset \mathcal{A}$, and, by 4.1.7, $\mu(\mathcal{A}) \subset \mu(\mathcal{F})$. ▷

4.1.12. A standard filter \mathcal{F} on X is an ultrafilter if and only if the monad of \mathcal{F} is easy to catch, i.e., for all standard subsets A and B in X such that $A \cup B = X$ either $\mu(\mathcal{F}) \subset A$ or $\mu(\mathcal{F}) \subset B$.

◁ \rightarrow : Since $\mu(\mathcal{F}) \subset A \cup B$, we may assume that $\mu(\mathcal{F}) \cap A \neq \emptyset$. Since $A = \mu(\{\text{fil } A\})$; therefore, $\mu(\mathcal{F}) \subset A$ by 4.1.11.

\leftarrow : Let $\mathcal{G} \supset \mathcal{F}$. Then, according to 4.1.7, $\mu(\mathcal{G}) \subset \mu(\mathcal{F})$. If A is standard and $A \supset \mu(\mathcal{G})$, then either $A \supset \mu(\mathcal{F})$ or $A' := X - A \supset \mu(\mathcal{F})$ by hypothesis. The case $A' \supset \mu(\mathcal{F})$ is impossible, since we would get $\mu(\mathcal{F}) \cap \mu(\mathcal{G}) \subset A \cap A' = \emptyset$. Therefore, $A \supset \mu(\mathcal{F})$ i.e., $A \in \mathcal{F}$ by 4.1.2. Consequently, $A \in \mathcal{F}$ for every standard $A \in \mathcal{G}$. By transfer, $\mathcal{G} \subset \mathcal{F}$, i.e., \mathcal{F} is an ultrafilter. ▷

4.1.13. A filter \mathcal{F} is an ultrafilter if and only if $A \cup B \in \mathcal{F} \rightarrow A \in \mathcal{F} \vee B \in \mathcal{F}$.

◁ \rightarrow : If $A \cup B \in \mathcal{F}$ then the monad of \mathcal{F} is caught: $\mu(\mathcal{F}) \subset A \cup B$. If $\mu(\mathcal{F}) \cap A \neq \emptyset$ then $\mu(\mathcal{F}) \subset A$ and $A \in \mathcal{F}$. If $\mu(\mathcal{F}) \cap B \neq \emptyset$ then $\mu(\mathcal{F}) \subset B$ and $B \in \mathcal{F}$.

\leftarrow : Let $A \cup B = X$. If $A \in \mathcal{F}$ then $A \supset \mu(\mathcal{F})$. If $B \in \mathcal{F}$ then $B \supset \mu(\mathcal{F})$; i.e., the monad is easily caught. ▷

4.1.14. Each limit of a filter is one of its adherent points. An adherent point of an ultrafilter is one of its limits.

◁ It suffices to assume the standard environment. Obviously, $\mathcal{F} \rightarrow x \leftrightarrow \mu(\mathcal{F}) \subset \mu(x) := \mu(\tau(x))$. Moreover, $x \in \text{cl}(\mathcal{F}) := \bigcap \{\text{cl}(F) : F \in \mathcal{F}\} \leftrightarrow (\forall F \in \mathcal{F})(\forall U \in \tau(x))(U \cap F \neq \emptyset) \leftrightarrow (\mu(\mathcal{F}) \cap \mu(x) \neq \emptyset)$ by 4.1.10 (1). We have thus proved the first claim.

Assume now that \mathcal{F} is an ultrafilter. Take $x \in \text{cl}(\mathcal{F})$. Then $\mu(\mathcal{F}) \cap \mu(x) \neq \emptyset$. Using the alternative of 4.1.11, deduce $\mu(\mathcal{F}) \subset \mu(x)$; i.e., $\mathcal{F} \rightarrow x$. ▷

4.1.15. Let \mathcal{E} be a cover of X . The following are equivalent:

- (1) There is a standard finite subcover \mathcal{E}_0 of \mathcal{E} , i.e., $\mathcal{E}_0 \in \mathcal{P}_{\text{st fin}}(\mathcal{E})$ and $X \subset \bigcup \mathcal{E}_0$;
- (2) The monad $\mu(\mathcal{E})$ of \mathcal{E} coincides with X ;
- (3) The monad $\mu(\mathcal{E})$ of \mathcal{E} is a standard set;
- (4) The monad $\mu(\mathcal{E})$ of \mathcal{E} is an internal set;
- (5) To each standard ultrafilter \mathcal{F} on X there is a member of \mathcal{E} belonging to \mathcal{F} .

◁ The implications (1) → (2) → (3) → (4) are all clear.

If $\mu(\mathcal{E})$ is an internal set then by 4.1.6(4) and 4.1.4 we conclude that $\mu(\mathcal{E})$ is standard; i.e., there is a standard finite $\mathcal{E}_0 \subset \mathcal{E}$ satisfying $\mu(\mathcal{E}) = \bigcup \mathcal{E}_0 \supset X$. Hence, (4) → (1).

The implication (1) → (5) is obvious. To prove (5) → (1), assume by way of contradiction that $(\forall^{\text{st fin}} \mathcal{E}_0) \cup \mathcal{E}_0 \not\subset X$. Put $\mathcal{E}' := \{E' := X - E : E \in \mathcal{E}\}$. The family \mathcal{E}' obviously generates a filterbase on X . Let \mathcal{F} be some ultrafilter finer than this filterbase. In this case we may find $E \in \mathcal{E}$ such that $E \in \mathcal{F}$. Moreover, by construction, $E' \in \mathcal{F}$. Thus, we come to a contradiction. ▷

4.1.16. In closing, we list some useful propositions that rest on the “technique of internal sets.”

4.1.17. Cauchy Principle. *Let \mathcal{F} be a standard filter on a standard set. Assume further that $\varphi := \varphi(x)$ is an internal property (i.e., $\varphi = \varphi^I$ for a set-theoretic formula φ). If $\varphi(x)$ holds for every remote element x relative to \mathcal{F} then there is a standard set $F \in \mathcal{F}$ such that $(\forall x \in F) \varphi(x)$.*

◁ There is an internal set F satisfying the property in question (for instance, a remote element of the filter \mathcal{F}). By transfer, we conclude that there is a standard set F enjoying the same property. ▷

4.1.18. Granted Horizon Principle. *Let X and Y be standard sets. Assume further that \mathcal{F} and \mathcal{G} are standard filters on X and Y respectively satisfying $\mu(\mathcal{F}) \cap {}^\circ X \neq \emptyset$. Distinguish a remote set, a “horizon,” F in ${}^a\mathcal{F}$. Given a standard correspondence $f \subset X \times Y$ meeting \mathcal{F} , the following are equivalent:*

- (1) $f(\mu(\mathcal{F}) - F) \subset \mu(\mathcal{G})$;
- (2) $(\forall F' \in {}^a\mathcal{F}) f(F' - F) \subset \mu(\mathcal{G})$;
- (3) $f(\mu(\mathcal{F})) \subset \mu(\mathcal{G})$.

◁ Obviously, (3) → (1) → (2). Hence, we have only to establish the implication (2) → (3).

To this end, choose a member G in ${}^\circ\mathcal{G}$. Suppose that for every standard F'' of ${}^\circ\mathcal{F}$ there is an element x of $F'' - F$ for which $f(x) \notin G$. By idealization, there is an element x' in $\mu(\mathcal{F})$ such that $x' \notin F$ and at the same time $f(x') \notin G$. We now put $F' := F \cup \{x'\}$. Clearly, $F' \in {}^a\mathcal{F}$, which leads to a contradiction implying that $f(F'' - F) \subset G$ for some standard $F'' \in \mathcal{F}$. Since F contains no standard elements of X , deduce

$$(\forall^{\text{st}} G \in \mathcal{G})(\exists^{\text{st}} F \in \mathcal{F})(\forall^{\text{st}} x \in F)(f(x) \in G).$$

It suffices to appeal to the transfer principle. ▷

4.2. Monads and Topological Spaces

In this section we study the properties of the monads of neighborhood filters in topological spaces.

4.2.1. Let (X, τ) be a standard *pretopological space*. Namely, to each standard point x of X there is a standard filter $\tau(x)$ on X of sets containing x . Put $\mu(x) := \mu_{\tau(x)} := \mu(\tau(x))$. We say that the elements of $\mu(x)$ are *infinitely close* to x . Obviously, $\mu(x)$ is the monad of the neighborhood filter $\tau(x)$ of x . A pretopological space (X, τ) is a *topological space* provided that each neighborhood of a point in X contains an open neighborhood of this point. In other words, each point x in ${}^\circ X$ has an *infinitely small* or *infinitesimal neighborhood* $U \in \tau(x)$ satisfying $\mu(x') \subset \mu(x)$ for all $x' \in U$.

4.2.2. Let G be an external set in a topological space (X, τ) . Put $h(G) := \bigcup \{\mu(x) : x \in {}^\circ G\}$. The set $h(G)$ is the *halo* of G in X . The set $G \cap h(G)$ is the *autohalo* or *nearstandard part* of G . The nearstandard part of G is denoted by $\text{nst}(G)$. If $G \supset h(G)$ then G is called *saturated* or, in more detail, *τ -saturated*. If for all $x \in G$ we have $\mu(x) \subset G$ then G is called *well-saturated* (*well- τ -saturated*).

4.2.3. *A standard set is open if and only if it is saturated.*

\triangleleft If G is open and $x \in {}^\circ G$ then $G \supset \mu(x)$. Hence, G includes its halo.

Conversely if $G \supset h(G)$ then choose a remote element U_x of the filter $\tau(x)$ for $x \in {}^\circ G$. Clearly, $G \supset U_x$. Consequently, G is open by transfer. \triangleright

4.2.4. A standard element x of X is a *microlimit* point of U provided that $\mu(x) \cap U \neq \emptyset$. We call the standardization of all microlimit points of U , the *microclosure* of U and denoted it by $\text{cl}_{\approx}(U)$.

4.2.5. *The microclosure $\text{cl}_{\approx}(U)$ of an arbitrary internal set U is closed. If U is a standard set then the microclosure $\text{cl}_{\approx}(U)$ coincides with the closure $\text{cl}(U)$ of U .*

\triangleleft Put $A := \text{cl}_{\approx}(U) = {}^*\{x \in X : \mu(x) \cap U \neq \emptyset\}$ and take $y \in \text{cl}(A)$. The task is to establish that $y \in A$. By transfer, it suffices to settle the case in which y is a standard point. Choose a standard open neighborhood V of y . By hypothesis, there is a standard point $x \in V$ such that $x \in A$. Using the definitions of standardization and monad, deduce that $V \supset \mu(x)$ and $\mu(x) \cap U \neq \emptyset$. Hence, $(\forall^{\text{st}} V \in \tau(y)) V \cap U \neq \emptyset$. By idealization, conclude that $\mu(y) \cap U \neq \emptyset$, i.e., $y \in \text{cl}_{\approx}(U)$.

Assume U standard. Obviously, ${}^\circ U \subset \text{cl}_{\approx}(U)$. By what was proved above, $U \subset \text{cl}_{\approx}(U)$ and $\text{cl}(U) \subset \text{cl}_{\approx}(U)$. Considering $y \in \text{cl}(U)$, observe that $(\forall^{\text{st}} V \in \tau(y)) V \cap U \neq \emptyset$. By idealization, $\mu(y) \cap U \neq \emptyset$, i.e., $y \in \text{cl}_{\approx}(U)$. \triangleright

4.2.6. For a point x and a nonempty set U the following are equivalent:

- (1) x is an adherent point of U ;
- (2) x is a microlimit point of U ;
- (3) There is a standard filter \mathcal{F} whose monad $\mu(\mathcal{F})$ lies in the monad $\mu(x)$;
- (4) There is a standard net $(x_\xi)_{\xi \in \Xi}$ of points of U whose entries with remote indices are infinitely close to x , i.e., $x_\xi \in \mu(x)$ for all $\xi \in {}^a\Xi$.

\triangleleft (1) \rightarrow (2): If $x \in \text{cl}(U)$ then the least upper bound $\tau(x) \vee \text{fil}\{U\}$ is available. Using 4.1.10(1), obtain

$$\emptyset \neq \mu(\tau(x) \vee \text{fil}\{U\}) = \mu(\tau(x)) \cap \mu(\text{fil}\{U\}) = \mu(x) \cap U,$$

which implies that $x \in \text{cl}_{\approx}(U)$.

(2) \rightarrow (3): If $x \in \text{cl}_{\approx}(U)$ then $U \cap \mu(x) \neq \emptyset$. Using 4.1.10(1), we may now construct a filter \mathcal{F} such that $A \in \mathcal{F} \leftrightarrow A \supset U \cap \mu(x)$. Obviously, such a filter \mathcal{F} gives what we seek.

(3) \rightarrow (4): Put $\Xi := \tau(x)$ and order Ξ as follows: $\xi_1 \leq \xi_2 \leftrightarrow \xi_1 \supset \xi_2$. Take as x_ξ an arbitrary point of $F \in \mathcal{F}$ such that $F \subset \xi$. Obviously, $(x_\xi)_{\xi \in \Xi}$ is a sought net. Indeed, $x_\xi \in \mu(x)$ for $\xi \in {}^a\Xi$ by construction.

(4) \rightarrow (1): Let V be a standard neighborhood of x , and let η be an arbitrary remote index of Ξ . Obviously, $x_\xi \in V$ for $\xi \geq \eta$, since $\mu(x) \subset V$ and $\xi \in {}^a\Xi$. It follows that $V \cap U \neq \emptyset$ since $x_\xi \in U$ by hypothesis. \triangleright

4.2.7. Assume that (X, τ) and (Y, σ) are standard topological spaces, $f : X \rightarrow Y$ is a standard mapping, and x is a standard point in X . The following are equivalent:

- (1) f is continuous at x ;
- (2) f sends each point infinitely close to x to a point infinitely close to $f(x)$, i.e.,

$$(\forall x')(x' \in \mu_\tau(x) \rightarrow f(x') \in \mu_\sigma(f(x))).$$

\triangleleft It suffices to refer to 4.1.8. \triangleright

4.2.8. Given a set A in X , we denote by $\mu(A)$ the intersection of standard open sets including A . The set $\mu(A)$ is the monad of A .

Note that $\mu(\emptyset) = \emptyset$. If $A \neq \emptyset$ then $\mu(A)$ is the monad of the neighborhood filter of A .

4.2.9. Let (X, τ) be a standard topological space. Then

- (1) (X, τ) is a separated ($= T_1$) space if and only if ${}^\circ\mu(x) = \{x\}$ for every point $x \in {}^\circ X$;
- (2) (X, τ) is a Hausdorff ($= T_2$) space if and only if $\mu(x_1) \cap \mu(x_2) = \emptyset$ for distinct $x_1, x_2 \in {}^\circ X$;

- (3) (X, τ) is regular if X is a T_1 -space enjoying the property T_3 : $\mu(x) \cap \mu(A) = \emptyset$ for every closed standard $A \subset X$ and every standard point $x \notin A$;
- (4) (X, τ) is normal if X is separated and satisfies the property T_4 : $\mu(A) \cap \mu(B) = \emptyset$ for every two disjoint closed sets A and B in X .

4.2.10. *The following are true:*

- (1) *A standard set is well-saturated if and only if it is open;*
- (2) *The monad of an arbitrary set is well-saturated;*
- (3) *The monad of a standard filter \mathcal{F} is well-saturated if and only if \mathcal{F} has a filterbase composed of open sets;*
- (4) *The monad $\mu(A)$ of an arbitrary internal A is the least well-saturated set including A ; moreover, $\mu(A) = \bigcup \{\mu(a) : a \in A\}$.*

\triangleleft (1): If A is standard and well-saturated then A is saturated and so A is open (see 4.2.3).

On assuming conversely that A is standard and open we apply the definition of monad to find $\mu(a) \subset A$ for $a \in A$, i.e., A is well-saturated.

(2): The monad of a set is by definition the intersection of standard open sets. Hence, this monad is well-saturated by (1).

(3): If \mathcal{F} has a base consisting of open standard sets then everything follows from (1). If $\mu(\mathcal{F})$ is well-saturated and V is a standard member of \mathcal{F} then $V \supset \mu(\mathcal{F}) \supset \bigcup \{U_a : a \in F\}$, where F is some remote element of \mathcal{F} , and U_a is an infinitesimal neighborhood of a . Since $\bigcup \{U_a : a \in F\} \in \mathcal{F}$, the claim follows by transfer.

(4): By (2), $\mu(A)$ is well-saturated. Moreover, so is the set $B := \bigcup \{\mu(a) : a \in A\}$ by (3). We are left with checking that $B = \mu(A)$. The inclusion $B \subset \mu(A)$ is obvious. Assume by way of contradiction that $B \neq \mu(A)$. Hence, there is some x in $\mu(A)$ satisfying $x \notin B$. Therefore, to each $a \in A$ there is a standard neighborhood U_a of a satisfying $x \notin U_a$. In other words, $(\forall a \in A)(\exists^{\text{st}} U_a) U_a \in \tau(a)$. By idealization, note that there is a standard finite set $\{a_1, \dots, a_n\} \subset A$ satisfying $A \subset U_{a_1} \cup \dots \cup U_{a_n}$. Hence, $x \in \mu(A) \subset U_{a_1} \cup \dots \cup U_{a_n}$, which is a contradiction. \triangleright

4.2.11. *Let (X, τ) be a separated topological space. A mapping $f : (X, \tau) \rightarrow (Y, \tau)$ is continuous at a point x if and only if $f(\mu_\tau(x) - U) \subset \mu_\sigma(f(x))$ for some (and hence every) infinitesimal neighborhood U of x .*

\triangleleft By separatedness $\mu_\tau(x) - U = \mu(x) - U$, where $\mu(x)$ is the monad of the filter $\tau(x)$ of deleted neighborhoods of x , i.e., $V \in \tau(x) \leftrightarrow V \cup \{x\} \in \tau(x)$. Obviously, $\mu(x) = \mu_\tau(x) - \{x\}$, in which case $U - \{x\}$ is an infinitesimal member of $\tau(x)$. Using the granted horizon principle 4.1.18, conclude that $f(\mu(x) - U) \subset \mu_\sigma(f(x)) \leftrightarrow f(\mu(x)) \subset \mu_\sigma(f(x)) \leftrightarrow f(\mu_\tau(x)) \subset \mu_\sigma(f(x))$. \triangleright

4.2.12. Let $(Y_\xi, \sigma_\xi)_{\xi \in \Xi}$ be a family of topological spaces. Assume further that $(f_\xi : X \rightarrow Y_\xi)_{\xi \in \Xi}$ is a family of mappings, and $\tau := \sup_{\xi \in \Xi} f_\xi^{-1}(\sigma_\xi)$ is the initial topology in X with respect to $(f_\xi : X \rightarrow Y_\xi)_{\xi \in \Xi}$, i.e., τ is the weakest topology in X such that the mapping f_ξ is continuous for all $\xi \in \Xi$. In this event,

$$\mu_\tau(x) = \bigcap_{\xi \in \circ\Xi} f_\xi^{-1}(\mu(\sigma_\xi(f_\xi(x))))$$

for every standard point $x \in X$.

◁ The claim is immediate from 4.1.8. ▷

4.2.13. A point x' of a Tychonoff product is infinitely close to a given point x provided that each standard coordinate of x' is close to the corresponding standard coordinate of x .

◁ Formally speaking, let $(X_\xi, \tau_\xi)_{\xi \in \Xi}$ be a standard family of standard topological spaces. Assume further that (\mathcal{X}, τ) is the Tychonoff product of $(X_\xi, \tau_\xi)_{\xi \in \Xi}$, i.e.,

$$\mathcal{X} := \prod_{\xi \in \Xi} X_\xi; \quad \tau := \sup_{\xi \in \Xi} \text{Pr}_\xi^{-1}(\tau_\xi),$$

where Pr_ξ is the projection from \mathcal{X} onto X_ξ . Given $x \in \circ\mathcal{X}$, we use 4.2.11 and 4.1.6 (5) to infer

$$\mu(x) = \bigcap_{\xi \in \circ\Xi} \mu(\text{Pr}_\xi^{-1}(\tau_\xi(x_\xi))) = \bigcap_{\xi \in \circ\Xi} \text{Pr}_\xi^{-1}(\mu(\tau_\xi(x_\xi))).$$

Given $\xi \in \circ\Xi$, observe $x' \in \text{Pr}_\xi^{-1}(\mu(\tau_\xi(x_\xi))) \leftrightarrow \text{Pr}_\xi x' \in \mu(\tau_\xi(x_\xi))$, i.e.,

$$\text{Pr}_\xi^{-1}(\mu(\tau_\xi(x_\xi))) = \mu_{\tau_\xi}(x_\xi) \times \prod_{\eta \neq \xi} X_\eta.$$

Therefore,

$$\text{Pr}_\xi(\mu(x)) = \mu(\tau_\xi(x_\xi))$$

for every standard $\xi \in \Xi$ (cf. 4.1.6 (5)), which completes the proof. ▷

4.3. Nearstandardness and Compactness

Proximity to a standard point, existent in topological spaces, makes it possible to give convenient tests for a set to be compact. These tests are the topic of the present section.

4.3.1. A point x of a topological standard space (X, τ) is *nearstandard* or, amply, τ -*nearstandard* provided that $x \in \text{nst}(X)$, i.e., if $x \in \mu(x')$ for some standard $x' \in {}^\circ X$.

4.3.2. A point $x \in X$ is nearstandard if and only if $x \in \mu(\mathcal{E})$ for each standard open cover \mathcal{E} of X . In symbols,

$$\text{nst}(X) = \bigcap \{ \mu(\mathcal{E}) : \mathcal{E} \text{ is an open cover of } X \}.$$

◁ Assume first that $x \in \mu(x')$ with $x \in \text{nst}(X)$ and $x' \in {}^\circ X$. Given an open cover \mathcal{E} , find a standard element $E \in \mathcal{E}$ satisfying $x' \in E$, i.e., $\mu(x') \subset E$ (see 4.2.3). Note that $x \in \mu(x') \subset E \subset \mu(\mathcal{E})$.

We now take $x \notin \text{nst}(X)$. We then have $x \notin \mu(x')$ for all $x' \in {}^\circ X$. Hence, there is a standard open neighborhood $U_{x'}$ of x' satisfying $x \notin U_{x'}$. The standardization $\mathcal{E} := {}^* \{ U_{x'} : x' \in {}^\circ X \}$ is an open cover of X such that $x \notin \mu(\mathcal{E})$. ▷

4.3.3. Each nearstandard point of X is infinitely close to a unique standard point if and only if X is Hausdorff.

◁ If τ is a Hausdorff topology and $x', x'' \in {}^\circ X$, then $\mu(x') \cap \mu(x'') \neq \emptyset \rightarrow x' = x''$. Conversely, assume that $x \in \mu(x') \cap \mu(x'')$ for $x', x'' \in {}^\circ X$. Since x is nearstandard; therefore, $x' = x''$ by hypothesis and so $x' \neq x'' \rightarrow \mu(x') \cap \mu(x'') = \emptyset$. ▷

4.3.4. Define the external correspondence $\text{st}(x) := \{ x' \in {}^\circ X : x \in \mu(x') \}$. In the Hausdorff case st is an external mapping from $\text{nst}(X)$ onto ${}^\circ X$.

4.3.5. If U is an internal set then $\text{cl}_\approx(U) = {}^* \text{st}(U)$. In particular, a standard set U is closed if and only if $U = {}^* \text{st}(U)$.

◁ Everything ensues from 4.2.5. ▷

4.3.6. Theorem. For a standard space X the following are equivalent:

- (1) X is compact;
- (2) Every point of X is nearstandard;
- (3) The autohalo of X is an internal set.

◁ (1) \rightarrow (2): Let \mathcal{E} be an open cover of X . The monad $\mu(\mathcal{E})$ coincides with X by 4.1.15 (since X is compact). By 4.3.2, $\text{nst}(X) = \bigcap_{\mathcal{E}} \mu(\mathcal{E}) = X$.

(2) \rightarrow (3): Obvious.

(3) \rightarrow (1): Let \mathcal{E} be a standard open cover of X . Since $(\forall x \in \text{nst}(X)) (\exists^{\text{st}} E \in \mathcal{E}) x \in E$; therefore, by idealization, $(\exists^{\text{st fin}} \mathcal{E}_0 \subset \mathcal{E}) \cup \mathcal{E}_0 \supset \text{nst}(X) \supset {}^\circ X$. By transfer, \mathcal{E}_0 is a cover of X . ▷

4.3.7. Let C be a set in a topological space X . The following are equivalent:

- (1) C is compact in the induced topology;
- (2) C lies in the halo $h(C)$;
- (3) The monad $\mu(C)$ coincides with the halo $h(C)$.

\triangleleft (1) \rightarrow (2): Since C is compact in the induced topology; therefore, $C \subset \text{nst}(C) \subset h(C)$ (see 4.3.6).

(2) \rightarrow (3): Clearly, $h(G) = \bigcup\{\mu(x) : x \in {}^\circ G\} \subset \mu(G)$. By hypothesis, to each $x \in C$ there is a member y in ${}^\circ C$ satisfying $x \in \mu(y)$. By 4.2.8 (2), $\mu(x) \subset \mu(y)$. By 4.2.8 (4), $\mu(C) = \bigcup\{\mu(x) : x \in C\} \subset \bigcup\{\mu(y) : y \in {}^\circ C\} = h(C)$.

(3) \rightarrow (1): Let \mathcal{E} be a standard open cover of C . By definition, $C \subset \mu(C) \subset h(C)$. Therefore (cf. 4.3.2), $C \subset \mu(\mathcal{E})$. By 4.1.15, there is a finite subcover of C in \mathcal{E} . \triangleright

4.3.8. For a regular space X and a set C in X the following are equivalent:

- (1) C is relatively compact (i.e., $\text{cl}(C)$ is compact);
- (2) C lies in the nearstandard part of X .

\triangleleft (1) \rightarrow (2): With no extra hypotheses, Proposition 4.3.7 obviously yields:

$$C \subset \text{cl}(C) \subset h(\text{cl}(C)) \subset h(X) = h(X) \cap X = \text{nst}(X).$$

(2) \rightarrow (1): Consider the closure $\text{cl}(C)$, and let \mathcal{E} be an open cover of $\text{cl}(C)$. Hence, to each $c \in C$ there is an $E \in \mathcal{E}$ containing c . Let E_c be a closed neighborhood of c included in E . Obviously, the family $\mathcal{E}' := \{E_c : c \in C\}$ is a standard cover of $\text{cl}(C)$. The family $\mathcal{E}' \cup \{X - \text{cl}(C)\}$ covers X and so, on using 4.3.1, we infer $C \subset \text{nst}(X) \subset \mu(\mathcal{E}' \cup \{X - \text{cl}(C)\})$. By 4.1.15, there is a finite set $\mathcal{E}_0 \subset \mathcal{E}'$ covering C . Obviously, $\bigcup \mathcal{E}_0$ is closed, i.e., \mathcal{E}_0 is a cover of $\text{cl}(C)$. Each element of \mathcal{E}_0 is, by construction, a subset of a member of \mathcal{E} . Therefore, we may refine a finite subcover of $\text{cl}(C)$ from the initial cover \mathcal{E} . \triangleright

4.3.9. The test for relative compactness 4.3.8 allows strengthening. Namely, the microclosure is compact of an arbitrary internal subset of the nearstandard part of an arbitrary Hausdorff space.

4.3.10. Let $\mathcal{X} := \prod_{\xi \in \Xi} X_\xi$ be the standard product of some standard family of topological spaces. A point $x \in \mathcal{X}$ is nearstandard if and only if so is its every standard coordinate $x_\xi \in \text{nst}(X_\xi)$ for $\xi \in {}^\circ \Xi$.

\triangleleft If $x \in \text{nst}(\mathcal{X})$ then $x_\xi \in \mu(y_\xi)$ for some $y \in {}^\circ \mathcal{X}$ and all $\xi \in {}^\circ \Xi$ by 4.1.12. It suffices to note that $y_\xi \in {}^\circ X_\xi$ by transfer.

Assume now that $x_\xi \in \text{nst}(X_\xi)$ for $\xi \in {}^\circ \Xi$. Consider the external function $y : \xi \mapsto \text{st}(x_\xi)$ from ${}^\circ \Xi$ to $\bigcup_{\xi \in \Xi} {}^\circ X_\xi$. Considering the standardization $*y$ and using 4.1.12, conclude easily that $*y \in {}^\circ \mathcal{X}$ and $x \in \mu(*y)$. \triangleright

4.3.11. Tychonoff Theorem. *The Tychonoff product of compact sets is compact.*

◁ By transfer, it suffices to settle the case of a standard family of standard spaces. In this event, appealing to 4.3.10, conclude that every point of the product is nearstandard. ▷

4.3.12. In sequel we usually consider Hausdorff compact spaces. It is in common parlance to refer to such a space briefly as a *compactum*.

4.4. Infinite Proximity in Uniform Space

Each uniform space generates a symmetric, reflexive, and transitive external relation between internal points, called *infinite proximity*. We will now address the most important relevant constructions.

4.4.1. Let (X, \mathcal{U}) be a uniform space. This implies that $\mathcal{U} := \{\emptyset\}$ in case $X = \emptyset$. If $X \neq \emptyset$ then \mathcal{U} is a filter on X^2 called the *uniformity* of X and enjoying the following properties:

- (1) $\mathcal{U} \subset \text{fil}\{I_X\}$;
- (2) $(\forall U \in \mathcal{U})(U^{-1} \in \mathcal{U})$;
- (3) $(\forall V \in \mathcal{U})(\exists U \in \mathcal{U})(U \circ U \subset V)$.

4.4.2. Luxemburg Test. *For a filter \mathcal{U} on X^2 to be a uniformity on a non-empty set X it is necessary and sufficient that the monad $\mu(\mathcal{U})$ of U is an external equivalence.*

◁ \rightarrow : Observe that

$$\begin{aligned} \mu(\mathcal{U}) &= \bigcap \circ\mathcal{U} = \bigcap_{U \in \circ\mathcal{U}} U = \bigcap_{U \in \circ\mathcal{U}} U^{-1} = \mu(\mathcal{U})^{-1}; \\ &\mu(\mathcal{U}) \supset I_X; \\ \mu(\mathcal{U}) &= \bigcap \{U \circ U : U \in \circ\mathcal{U}\} \supset \mu(\mathcal{U}) \circ \mu(\mathcal{U}) \supset \mu(\mathcal{U}) \circ I_X \supset \mu(\mathcal{U}). \end{aligned}$$

The above derivation rests on the fact that U^{-1} and $U \circ U$ are standard whenever so is U . Moreover, by the definition of monad, $U \supset \mu(\mathcal{U})$ for $U \in \circ\mathcal{U}$.

\leftarrow : By 4.1.4, the filter \mathcal{U} is the standardization of all supersets of the monad of \mathcal{U} , i.e.,

$$U \in \circ\mathcal{U} \leftrightarrow U \supset \mu(\mathcal{U}).$$

This implies that $\mathcal{U} \subset \text{fil}\{I_X\}$ and $U \in \mathcal{U} \rightarrow U^{-1} \in \mathcal{U}$. Consider an infinitesimal member W of the filter \mathcal{U} . By above, $U := W^{-1} \cap W \in \mathcal{U}$. Moreover, $U \circ U \subset \mu(\mathcal{U}) \circ \mu(\mathcal{U}) = \mu(\mathcal{U})$. Hence, to each standard $V \in \mathcal{U}$, there is some U in \mathcal{U} satisfying $U \circ U \subset V$. By transfer, conclude that \mathcal{U} is a uniformity. ▷

4.4.3. Using the Luxemburg test it is worthwhile to bear in mind that an arbitrary external equivalence on X^2 may fail to be the monad of any filter (generating no uniformity on X in this event). For instance, assuming that $x, y \in \mathbb{R}$ are equivalent whenever $x - y \in {}^\circ\mathbb{R}$, we see the external coset of zero is the standard part ${}^\circ\mathbb{R}$ of the reals \mathbb{R} which is the monad of no filter. We conclude in particular that this external equivalence produces no standard uniformity.

4.4.4. Consider a couple of points x and y in a space X with uniformity \mathcal{U} . The points x and y are *infinitely close* (relative to \mathcal{U}), in symbols $x \approx_{\mathcal{U}} y$ or $x \approx y$, provided that $(x, y) \in \mu(\mathcal{U})$. If A is an arbitrary (possibly, external) set in X then we call the external set $\mu_{\mathcal{U}}(A)$ the *microhalo* of A in X and denoted it by $\approx A$. If A is a standard set then, slightly abusing consistency, we let the symbol $\approx A$ denote the halo $h(A)$ of A , implying the equality $h(A) = \approx^\circ A$. Clearly, in this event we calculate the halo relative to the uniform topology $\tau_{\mathcal{U}}$ on X generated by \mathcal{U} . Observe that in this topology the monad of a standard point x consists, as it might be expected, of all points infinitely close to x , i.e., the monad of x coincides with the microhalo $\approx x := \approx\{x\}$ of x . A less adequate terminology is used sometimes in which the microhalo $\approx x$ of an internal point x is still called the monad of x . This is misleading since the microhalo of a point may fail to be the monad of any filter.

4.4.5. A function f from a uniform space X to a uniform space Y is *microcontinuous* on X provided that f sends infinitely close points of X to infinitely close points of Y .

4.4.6. Theorem. *The following hold:*

- (1) *A standard function is microcontinuous if and only if it is uniformly continuous;*
- (2) *A standard set consists of microcontinuous functions if and only if it is a (uniformly) equicontinuous set.*

\triangleleft (1): The uniform continuity of $f : X \rightarrow Y$ implies that $f^\times(\mathcal{U}_X) \supset \mathcal{U}_Y$, with \mathcal{U}_X and \mathcal{U}_Y standing for the uniformities of X and Y respectively and $f^\times(x, x') := (f(x), f(x'))$ for $x, x' \in X$. Considering 4.1.8, deduce

$$f^\times(\mathcal{U}_X) \supset \mathcal{U}_Y \leftrightarrow \mu(f^\times(\mathcal{U}_X)) \subset \mu(\mathcal{U}_Y).$$

(2): Recall that $\mathcal{E} \subset Y^X$ is a (uniformly) equicontinuous set whenever

$$(\forall V \in \mathcal{U}_Y)(\exists U \in \mathcal{U}_X)(\forall f \in \mathcal{E})(f^{-1} \circ V \circ f \in \mathcal{U}_X).$$

Given such a set \mathcal{E} , by transfer we have

$$(\forall^{\text{st}} V \in \mathcal{U}_Y)(\exists^{\text{st}} U \in \mathcal{U}_X)(\forall f \in \mathcal{E})(\forall x, x' \in U)((f(x), f(x')) \in V).$$

In particular, if $x \approx x'$ then $(f(x), f(x')) \in V$ for all $f \in \mathcal{E}$ and $V \in {}^\circ\mathcal{U}_Y$, i.e., $f(x) \approx f(x')$. Thus, an equicontinuous standard set contains only microcontinuous elements.

To prove the reverse implication, we will use the Cauchy principle 4.1.17 for the sake of diversity.

Indeed, given $V \in {}^\circ\mathcal{U}_Y$ and an arbitrary remote element $U \in {}^\circ\mathcal{U}_X$, note that $(\forall f \in \mathcal{E}) f^\times(U) \subset V$. Hence, there is some standard $U \in \mathcal{U}_X$ enjoying the same internal property. We are left with appealing to the transfer principle. \triangleright

4.4.7. Let (X, \mathcal{U}_X) and (Y, \mathcal{U}_Y) be standard uniform spaces, and let f be an internal function; $f : X \rightarrow Y$. Let ${}^E\mathcal{U}_X$ and ${}^E\mathcal{U}_Y$ stand for the filters of external supersets of ${}^\circ\mathcal{U}_X$ and ${}^\circ\mathcal{U}_Y$, respectively. In this case the following are equivalent:

- (1) f is microcontinuous;
- (2) $f : (X, {}^E\mathcal{U}_X) \rightarrow (Y, {}^E\mathcal{U}_Y)$ is uniformly continuous;
- (3) $(\forall {}^{\text{st}}V \in \mathcal{U}_Y)(\exists {}^{\text{st}}U \in \mathcal{U}_X)(f^\times(U) \subset V)$.

\triangleleft (1) \rightarrow (3): Take $V \in {}^\circ\mathcal{U}_Y$. Considering a remote element $U \in {}^\circ\mathcal{U}_X$, note that $(x, x') \in U \rightarrow x \approx x' \rightarrow f(x) \approx f(x')$, i.e., $f^\times(U) \subset V$. By the Cauchy principle 4.1.17, there is a standard set U enjoying the same property.

(3) \rightarrow (1): Take $x \approx x'$ and a standard element $V \in \mathcal{U}_Y$. By hypothesis, there is some standard $U \in \mathcal{U}_X$ satisfying $f^\times(U) \subset V$. In particular, $(f(x), f(x')) \in V$ and so $f(x) \approx f(x')$.

(3) \leftrightarrow (2): Obvious. \triangleright

4.4.8. EXAMPLES.

(1) Let X be a set, and let d be a *semimetric* (= *deviation*) on X . In other words, assume given the (standard) objects X and $d : X^2 \rightarrow \mathbb{R}$ such that

$$\begin{aligned} d(x, x) &= 0 \quad (x \in X); \\ d(x, y) &= d(y, x) \quad (x, y \in X); \\ d(x, y) &\leq d(x, z) + d(z, y) \quad (x, y, z \in X). \end{aligned}$$

Consider the cylinders $\{d \leq \varepsilon\} := \{(x, y) \in X^2 : d(x, y) \leq \varepsilon\}$ and the family $\mathcal{U}_d := \text{fil} \{\{d \leq \varepsilon\} : \varepsilon \in \mathbb{R}, \varepsilon > 0\}$. Obviously, \mathcal{U}_d furnishes X with the structure of a uniform space, namely, the semimetric uniformity of the *semimetric space* (X, d) . It is worth observing that the monad of the semimetric uniformity amounts to the following external equivalence:

$$x \approx_d y \leftrightarrow d(x, y) \approx 0 \leftrightarrow d(x, y) \in \mu(\mathbb{R}).$$

(2) Let (X, \mathfrak{M}) be a *multimetric space*, i.e., \mathfrak{M} is a *multimetric* (= a nonempty set of semimetrics on X). The monad $\mu(\mathfrak{M})$ of \mathfrak{M} is defined as the

intersection of the monads of the (standard) uniform spaces (X, d) , with $d \in {}^\circ\mathfrak{M}$. Namely,

$$x \approx_{\mathfrak{M}} y \leftrightarrow (\forall d \in {}^\circ\mathfrak{M}) d(x, y) \approx 0.$$

The monad $\mu(\mathfrak{M})$ is undoubtedly the monad of the uniformity $\mathcal{U}_{\mathfrak{M}} := \sup\{\mathcal{U}_d : d \in \mathfrak{M}\}$ of the multimetric space (X, \mathfrak{M}) . It stands to reason to recall that every uniform space (X, \mathcal{U}) is multimetrizable, i.e., $\mathcal{U} = \mathcal{U}_{\mathfrak{M}}$ for some multimetric \mathfrak{M} .

(3) Let (X, \mathcal{U}) be a uniform space. Furnish the powerset $\mathcal{P}(X)$ with the *Vietoris uniformity* whose neighborhood filterbase comprises the sets

$$\{(A, B) \in \mathcal{P}(X)^2 : B \subset U(A), A \subset U(B)\},$$

where $U \in \mathcal{U}$. Obviously, the monad $\mu_v := \mu_v(\mathcal{U})$ of the Vietoris uniformity is as follows:

$$\mu_v = \{(A, B) : A \subset \approx B, B \subset \approx A\}.$$

(4) Let (X, τ) be a compactum, i.e., a Hausdorff compact space. The space X is uniquely uniformizable, i.e., there is a unique uniformity \mathcal{U} on X^2 such that the uniform topology $\tau_{\mathcal{U}}$ coincides with the original topology τ . This uniformity \mathcal{U} is simply the neighborhood filter of the diagonal of X^2 . Therefore, $\mu(\mathcal{U}) = \mu_{\tau \times \tau}(I_X)$. In other words, $x \approx x' \leftrightarrow \text{st}(x) = \text{st}(x')$, since $\mu_{\tau \times \tau}(x, x) = \mu_{\tau}(x) \times \mu_{\tau}(x)$ for a standard point x (see 4.2.13) and every point of X^2 is near-standard (see 4.3.6).

(5) Let X and Y be nonempty sets, let \mathcal{U}_Y be a uniformity on Y , and let \mathcal{B} be a family of subsets of X which is upward-filtered by inclusion. Consider the uniformity \mathcal{U} on Y^X , called the “*uniformity of uniform convergence on the members of \mathcal{B}* .” The family \mathcal{U} is the union of supersets of the following elements:

$$V_{B,U} := \{(f, g) \in Y^X \times Y^X : g \circ I_B \circ f^{-1} \subset U\},$$

where $B \in \mathcal{B}$ and $U \in \mathcal{U}_Y$. It is immediate that

$$\begin{aligned} (f, g) \in \mu(\mathcal{U}) &\leftrightarrow (\forall^{\text{st}} B \in \mathcal{B})(\forall^{\text{st}} U \in U_Y)(\forall x \in B)(f(x), g(x) \in U) \\ &\leftrightarrow (\forall^{\text{st}} B \in \mathcal{B})(\forall x \in B)(f(x) \approx g(x)) \leftrightarrow (\forall x \in \mu(\mathcal{B}))(f(x) \approx g(x)), \end{aligned}$$

where, as usual, $\mu(\mathcal{B}) := \bigcup^{\circ} \mathcal{B}$ is the monad of \mathcal{B} . If $\mathcal{B} = \{X\}$ then we speak about the *strong uniformity* \mathcal{U}_s on X . The following relation is obvious:

$$(f, g) \in \mu(\mathcal{U}_s) \leftrightarrow (\forall x \in X)(f(x) \approx g(x)).$$

If $\mathcal{B} = \mathcal{P}_{\text{fin}}(X)$ then $\mu(\mathcal{B}) = {}^\circ X$ and so

$$(f, g) \in \mu(\mathcal{U}_w) \leftrightarrow (\forall^{\text{st}} x \in X)(f(x) \approx g(x))$$

for the corresponding *weak convergence* \mathcal{U}_w (or, which is the same by definition, for the *uniformity of pointwise convergence*).

4.4.9. A set A is *infinitely small* or *infinitesimal* (relative to the uniformity \mathcal{U}) provided that $A^2 \subset \mu(\mathcal{U})$, i.e., if any two points of A are infinitely close.

4.4.10. For a standard filter \mathcal{F} on (X, \mathcal{U}) the following hold:

- (1) The monad $\mu(\mathcal{F})$ is infinitesimal;
- (2) \mathcal{F} is a Cauchy filter;
- (3) For every $U \in {}^\circ\mathcal{U}$ there is some member x of ${}^\circ X$ such that $\mu(\mathcal{F}) \subset U(x)$.

\triangleleft (1) \rightarrow (2): Suppose that $\mu(\mathcal{F})^2 \subset \mu(\mathcal{U})$. Obviously $\mu(\mathcal{F})^2 = \mu(\mathcal{F}^\times)$, where $\mathcal{F}^\times := \{F^2 : F \in \mathcal{F}\}$, since

$$(x, y) \in \mu(\mathcal{F}^\times) \leftrightarrow (\forall^{\text{st}} F \in \mathcal{F})(x \in F \wedge y \in F) \leftrightarrow (x \in \mu(\mathcal{F}) \wedge y \in \mu(\mathcal{F})).$$

Therefore, $\mu(\mathcal{F}^\times) \subset \mu(\mathcal{U})$, i.e., $\mathcal{F}^\times \supset \mathcal{U}$. This implies that \mathcal{F} is a Cauchy filter.

(2) \rightarrow (3): For $U \in {}^\circ\mathcal{U}$ there is a standard element $F \in \mathcal{F}$ satisfying $F \times F \subset U$. If $x \in {}^\circ F$ then $(\forall^{\text{st}} y \in F)(y \in U(x))$. Hence, $F \subset U(x)$ and, moreover, $\mu(\mathcal{F}) \subset U(x)$.

(3) \rightarrow (1): By standardization, $(\exists x \in X) \mu(\mathcal{F}) \subset \approx x$. Hence, $\mu(\mathcal{F})$ is an infinitesimal external set. \triangleright

4.4.11. A Cauchy filter converges if and only if its monad contains a nearstandard point.

$\triangleleft \rightarrow$: If \mathcal{F} is a convergent Cauchy filter then $\mu(\mathcal{F}) \subset \mu(x)$ since $\mathcal{F} \rightarrow x$. Hence, every point of $\mu(\mathcal{F})$ is nearstandard.

\leftarrow : Let $\mu(\mathcal{F}) \cap \approx x \neq \emptyset$. Given $y \in \mu(\mathcal{F})$ and $z \in \mu(\mathcal{F}) \cap \approx x$, observe $y \approx z \approx x$, i.e., $y \approx x$. Hence, $\mu(\mathcal{F}) \subset \mu(x)$. It suffices to appeal to 4.1.7. \triangleright

4.5. Prestandardness, Compactness, and Total Boundedness

The uniform spaces are well-known to enjoy a convenient test for compactness, the classical Hausdorff test. In this section we give its nonstandard analogs in the standard environment as well as a relevant test for a point to be prestandard in the space of continuous functions on a compact space.

4.5.1. For an internal point x of a standard uniform space X the following are equivalent:

- (1) The microhalo of x is the monad of some standard filter on X ;
- (2) The microhalo of x is the monad of some Cauchy filter on X ;
- (3) The microhalo of x coincides with the monad of an inclusion-minimal Cauchy filter;
- (4) The microhalo of x contains some infinitesimal monad;
- (5) There is a standard net $(x_\xi)_{\xi \in \Xi}$ in X elements microconvergent to x , i.e., such that $x_\xi \approx x$ for all $\xi \in {}^a\Xi$.

◁ (1) → (2): If $\approx x = \mu(\mathcal{F})$ for some standard filter \mathcal{F} then $\mu(\mathcal{F})$ is infinitesimal (since so is the microhalo $\approx x$).

(2) → (3): Suppose that $\approx x = \mu(\mathcal{F})$ while \mathcal{F}' is a Cauchy filter and $\mathcal{F}' \subset \mathcal{F}$. Then $\mu(\mathcal{F}') \supset \mu(\mathcal{F}) = \approx x$ by 4.1.17. If $y \in \mu(\mathcal{F}')$ then $y \approx x$ since $\mu(\mathcal{F}')$ is infinitesimal, and so $\mu(\mathcal{F}') = \mu(x) = \mu(\mathcal{F})$. Hence, $\mathcal{F}' = \mathcal{F}$ (see 4.1.4).

(3) → (4): Obvious.

(4) → (1): Assume that $\approx x \supset \mu(\mathcal{F})$ and \mathcal{F} is a Cauchy filter. Put $\mathcal{F}' := \text{fil}\{U(F) : U \in \mathcal{U}_X, F \in \mathcal{F}\}$. Considering $\mathcal{U} := \mathcal{U}_X$, find

$$\begin{aligned} \approx \mu(\mathcal{F}) &= \mu(\mathcal{U})(\mu(\mathcal{F})) = \mu(\mathcal{U})\left(\bigcap_{F \in \circ \mathcal{F}} F\right) = \bigcap_{F \in \circ \mathcal{F}} \mu(\mathcal{U})(F) \\ &= \bigcap_{F \in \circ \mathcal{F}} \bigcap_{U \in \circ \mathcal{U}} U(F) = \bigcap \{F' : F' \in \circ \mathcal{F}'\} = \mu(\mathcal{F}'). \end{aligned}$$

Clearly, $\approx \mu(\mathcal{F}) \supset \approx x$, and so $\mu(\mathcal{F}) = \approx x = \mu(\mathcal{F}')$.

(4) → (5): If \mathcal{F} is a filter and $\mu(\mathcal{F}) \subset \approx x$ then, choosing a point from each standard $F \in \circ \mathcal{F}$ in a routine manner and applying standardization, construct a sought net. Conversely, if $(x_\xi)_{\xi \in \Xi}$ microconverges to x then the monad of the tail filter of this sequence lies in the microhalo $\approx x$. ▷

4.5.2. A point x , obeying one (and hence all) of the equivalent conditions 4.5.1 (1)–(4), is *prenearstandard* in X . The external set of prenearstandard points in X is denoted by $\text{pst}(X)$.

4.5.3. *Every nearstandard point relative to a uniform topology is prenearstandard.*

◁ Take $x \in \text{nst}(X)$, with (X, \mathcal{U}) the space under study. Therefore, $x \in \approx y$ for some $y \in \circ X$. Hence, $\approx x \supset \approx y = \mu(\tau_{\mathcal{U}}(y))$. By 4.5.1 conclude that $x \in \text{pst}(X)$. ▷

4.5.4. *The image of a prenearstandard point under a uniformly continuous mapping is prenearstandard.*

◁ Let \mathcal{F} be a Cauchy filter and $\mu(\mathcal{F}) \subset \approx x$. Obviously, $f(\mathcal{F})$ is a Cauchy filter on the image of X under the mapping f . Hence, $\mu(f(\mathcal{F})) \subset \approx f(x)$; i.e., $f(x)$ is a prenearstandard point (see 4.5.2). ▷

4.5.5. *A point of the Tychonoff product of a standard family of uniform spaces is prenearstandard if and only if so are its standard coordinates.*

◁ →: Let $(\mathcal{X}, \mathcal{U}_{\mathcal{X}})$ be the Tychonoff product of standard spaces $(X_\xi, \mathcal{U}_\xi)_{\xi \in \Xi}$; i.e., $\mathcal{X} = \prod_{\xi \in \Xi} X_\xi$ and $\mathcal{U}_{\mathcal{X}} = \sup_{\xi \in \Xi} \text{Pr}_\xi^{x^{-1}}(\mathcal{U}_\xi)$. Take $x \in \text{pst}(\mathcal{X})$. By 4.5.1, there is a Cauchy filter \mathcal{F} on $(\mathcal{X}, \mathcal{U}_{\mathcal{X}})$ satisfying $\approx x = \mu(\mathcal{F})$.

Since Pr_ξ is continuous; therefore, by 4.4.6, given a standard $\xi \in \Xi$, observe that $\text{Pr}_\xi(\approx x) \subset \approx x_\xi$, i.e., $\approx x_\xi \supset \text{Pr}_\xi(\mu(\mathcal{F})) = \mu(\text{Pr}_\xi(\mathcal{F}))$. Consequently, x_ξ is a prenearstandard point in X_ξ for all $\xi \in {}^\circ\Xi$.

\leftarrow : Assume that to each $\xi \in {}^\circ\Xi$ there is some Cauchy filter \mathcal{F}_ξ such that $\approx x_\xi = \mu(\mathcal{F}_\xi)$. Consider the filter

$$\mathcal{F} := \sup_{\xi \in \Xi} \text{Pr}_\xi^{-1}(\mathcal{F}_\xi).$$

Obviously, \mathcal{F} is a standard filter on \mathcal{X} and

$$\begin{aligned} \mu(\mathcal{F}) &= \bigcap_{\xi \in {}^\circ\Xi} \mu(\text{Pr}_\xi^{-1}(\mathcal{F}_\xi)) = \bigcap_{\xi \in {}^\circ\Xi} \text{Pr}_\xi^{-1}(\mu(\mathcal{F}_\xi)) \\ &= \bigcap_{\xi \in {}^\circ\Xi} \text{Pr}_\xi^{-1}(\approx x_\xi) = \{y \in \mathcal{X} : (\forall \xi \in {}^\circ\Xi) y_\xi \approx x_\xi\} = \approx x. \end{aligned}$$

The proof is complete. \triangleright

4.5.6. *A standard space is complete if and only if its every prenearstandard point is nearstandard.*

$\triangleleft \rightarrow$: Since X is complete, every Cauchy filter on X converges. Take $x \in \text{pst}(X)$. By 4.5.2, $\mu(\mathcal{F}) = \approx x$ for some Cauchy filter \mathcal{F} . By compactness, there is a point y in ${}^\circ X$ such that $\mu(y) \supset \mu(\mathcal{F})$. Hence, $\approx y = \mu(y) \supset \mu(\mathcal{F}) \supset \approx x$. Therefore, $\approx y = \approx x$, i.e., $x \in \text{nst}(X)$.

\leftarrow : Assume that $\text{nst}(X) = \text{pst}(X)$ and let \mathcal{F} be a Cauchy filter on X . Take a point x in $\mu(\mathcal{F})$. Then $\approx x \supset \mu(\mathcal{F})$ (since $\mu(\mathcal{F})$ is an infinitesimal set). By 4.5.2, $x \in \text{pst}(X)$. Hence, $x \in \text{nst}(X)$. We are done on appealing to 4.4.11. \triangleright

4.5.7. *The Tychonoff product of complete uniform spaces is complete.*

\triangleleft By transfer, assume the standard environment. Every prenearstandard point of a complete standard factor is nearstandard by 4.5.5. Recall now that a nearstandard point is a point with nearstandard standard coordinates by 4.3.10, while a prenearstandard point is a point with prenearstandard standard coordinates by 4.5.5. The proof is complete on chanting: “The uniform topology of a product is the product of the uniform topologies of factors.” \triangleright

4.5.8. *The space of functions acting into a complete space is complete under the strong uniformity.*

\triangleleft Let (Y, \mathcal{U}) be a complete standard uniform space, and let X be a standard set. Choose a prenearstandard point $f \in Y^X$. By 4.5.2 and 4.4.8, there is a standard net $(f_\xi)_{\xi \in \Xi}$ of elements of Y^X such that

$$(\forall \xi \in {}^a\Xi)(\forall x \in X)(f_\xi(x) \approx f(x)).$$

By 4.5.7, f is nearstandard in the weak uniformity, i.e., there is a standard element $g \in Y^X$ satisfying

$$(\forall \xi \in {}^a\Xi)(\forall {}^{\text{st}}x \in X)(f_\xi(x) \approx g(x)).$$

Hence, for every standard $x \in X$ the sequence $(f_\xi(x))_{\xi \in \Xi}$ converges to $g(x)$.

By transfer, $(\forall x \in X) f_\xi(x) \rightarrow g(x)$. Consequently,

$$(\forall U \in {}^\circ\mathcal{U})(\forall x \in X)(f(x), g(x)) \in U,$$

which implies that f is infinitely close to g in the strong uniformity. The proof is complete by referring to 4.5.6 and the transfer principle. \triangleright

4.5.9. *Let E be a set in a uniform space (X, \mathcal{U}) . The following are equivalent:*

- (1) E is a totally bounded set; i.e., to each $U \in \mathcal{U}$ there is a finite set $E_0 \subset E$ such that $E \subset U(E_0)$ (in other words, E has a finite U -net for all $U \in \mathcal{U}$);
- (2) there is an internal finite cover of E by infinitesimal internal sets;
- (3) E has a finite skeleton; i.e., there is an internal finite set E_0 in X such that E lies in the microhalo $\approx E_0$;
- (4) E lies in the microhalo of an internal totally bounded set.

\triangleleft (1) \leftrightarrow (2): By definition and idealization, infer successively that

$$\begin{aligned} & (\forall {}^{\text{st}}U \in \mathcal{U})(\exists E_0)(E_0 \subset E \wedge E_0 \in \mathcal{P}_{\text{fin}}(X) \wedge E \subset U(E_0)) \\ \leftrightarrow & (\forall {}^{\text{st}} \text{fin} \mathcal{U}_0 \subset \mathcal{U})(\exists E_0)(\forall U \in \mathcal{U}_0)(E_0 \subset E \wedge E_0 \in \mathcal{P}_{\text{fin}}(X) \wedge E \subset U(E_0)) \\ \leftrightarrow & (\exists E_0)(\forall {}^{\text{st}}U \in \mathcal{U})(E_0 \subset E \wedge E_0 \in \mathcal{P}_{\text{fin}}(X) \wedge E \subset U(E_0)) \\ \leftrightarrow & (\exists E_0 \subset E)(E_0 \in \mathcal{P}_{\text{fin}}(X) \wedge E \subset \approx E_0). \end{aligned}$$

(1) \leftrightarrow (3): Obviously, E is totally bounded if and only if to each standard $U \in \mathcal{U}$ there is a finite cover $\{E_1, \dots, E_n\}$ of E such that $E_k \times E_k \subset U$ (i.e., E_k is small of order U) for $k := 1, \dots, n$. We are done on appealing to the idealization principle.

(3) \rightarrow (4): Obvious.

(4) \rightarrow (1): Let U be a standard entourage of the diagonal. There is a symmetric element $V \in {}^\circ\mathcal{U}$ satisfying $V \circ V \subset U$. Given a finite set E' in X , we easily see that $V(E') \supset E_0$, where E_0 is a given totally bounded set enjoying the property $\approx E_0 \supset E$. Hence, $U(E') \supset V \circ V(E') \supset V(E_0) \supset E$. \triangleright

4.5.10. *Each standard uniform space X has a universal finite skeleton, i.e. a common internal finite skeleton for all totally bounded standard sets of X .*

\triangleleft Recall that the union of finitely many totally bounded sets is totally bounded. Given a finite standard family \mathcal{E} of totally bounded sets in X and a standard finite family $\mathcal{U}_0 \subset \mathcal{U}_X$, by 4.5.9, find a common finite set in X serving as a U -net for all $E \in \mathcal{E}$ and $U \in \mathcal{U}_0$. Proceed by idealization. \triangleright

4.5.11. For a uniform space X the following are equivalent:

- (1) X is totally bounded;
- (2) Every point of X is prenearstandard;
- (3) $\text{pst}(X)$ is an internal set;
- (4) X has a finite skeleton.

\triangleleft (1) \rightarrow (2): Take $x \in X$. To each standard $U \in \mathcal{U}$ there is a standard point $x' \in {}^\circ X$ such that $x \in U(x')$ is an element of a finite standard U -net for X . Put $\mathcal{F} := \text{fil}^* \{U(x') : U \in {}^\circ \mathcal{U}\}$. Obviously, \mathcal{F} is a Cauchy filter (see 4.4.10). In this event, $x \in \mu(\mathcal{F})$ by construction, i.e., $x \in \text{pst}(X)$.

(2) \rightarrow (3): Obvious.

(3) \rightarrow (1): Assume that the inclusion $\text{pst}(X) \subset U(E)$ is false for some standard $U \in \mathcal{U}$ and all finite standard subsets E of X . By idealization, there is an internal point $x \in \text{pst}(X)$ satisfying $x \notin U(y)$ for all $y \in {}^\circ X$. By 4.5.2, $\approx x = \mu(\mathcal{F})$ for some Cauchy filter \mathcal{F} . Choose $F \in {}^\circ \mathcal{F}$ so that $F \times F \subset U$. For all $y \in {}^\circ F$ we then see that $x \in \mu(\mathcal{F}) \subset U(y)$, which is a contradiction. Therefore, $(\forall^{\text{st}} U \in \mathcal{U})(\exists^{\text{st fin}} E \subset X) U(E) \supset \text{pst}(X)$. It suffices to recall that $\text{pst}(X) \supset {}^\circ X$.

(1) \leftrightarrow (4): Follows from 4.5.9. \triangleright

4.5.12. Hausdorff Test. A uniform space is compact if and only if it is complete and totally bounded.

$\triangleleft \rightarrow$: If X is a compact and standard space then every point of X is nearstandard and, hence, prenearstandard (by 4.5.3). By 4.5.11, X is totally bounded and X is complete by 4.5.6.

\leftarrow : Since X is totally bounded; therefore, $X = \text{pst}(X)$ by 4.5.11. Since X is complete, $\text{pst}(X) = \text{nst}(X)$ by 4.5.6. All in all, $X = \text{nst}(X)$; i.e., X is compact by 4.3.6. \triangleright

4.5.13. Assume that X is an arbitrary set, Y is a uniform space, and $f : X \rightarrow Y$ is a function. The following are equivalent:

- (1) f is a totally bounded mapping, i.e., $\text{im}(f)$ is totally bounded in Y ;
- (2) There is an internal finite cover \mathcal{E} of X such that $f(E)$ is infinitesimal for all $E \in \mathcal{E}$, i.e., f is a nearstep function relative to \mathcal{E} ;
- (3) There are an internal $n \in \mathbb{N}$ and a set $\{X_1, \dots, X_n\}$ of disjoint external sets such that $X_1 \cup \dots \cup X_n = X$ and $f(x) \approx (x')$ for all $x, x' = X_k$ and $k := 1, \dots, n$.

\triangleleft (1) \rightarrow (2): By 4.5.9, there is an internal finite cover \mathcal{E} of $\text{im}(f)$ such that $E \in \mathcal{E} \rightarrow E_2 \subset \mu(\mathcal{U}_Y)$. Put $\mathcal{E}' := \{f^{-1}(E) : E \in \mathcal{E}\}$. Obviously, \mathcal{E}' is a sought cover of X .

(2) \rightarrow (3): Obvious.

(3) \rightarrow (1): Choose $y_k \in f(X_k)$ and put $E := \{y_k : k = 1, \dots, n\}$. Clearly, E is a finite internal set. By hypothesis, E is a skeleton of $f(X)$. Hence, $\text{im}(f)$ is totally bounded by 4.5.9. \triangleright

4.5.14. *The space $CB(X, Y)$ of totally bounded mappings from X to Y is complete in the strong uniformity.*

\triangleleft By 4.5.8, it suffices to demonstrate that $CB(X, Y)$ is closed. To this end, let a standard $f : X \rightarrow Y$ be such that $(\forall x \in X) f(x) \approx g(x)$ for some totally bounded function g . Clearly, $f \subset \approx \text{im}(g)$. Since $\text{im}(g)$ is totally bounded, from 4.2.5 and 4.5.9 infer that $f \in \text{cl}(CB(X, Y)) \rightarrow f \in CB(X, Y)$. \triangleright

4.5.15. A finite cover \mathcal{E} of a standard set X is *tiny* provided that \mathcal{E} coarsens every standard finite cover \mathcal{E}_0 of X , i.e., if every member of \mathcal{E} is included in some member of \mathcal{E}_0 . A mapping f from X into a uniform space is *microstep* on X provided that f is nearstep relative to every tiny cover of X .

4.5.16. *A function $f : X \rightarrow Y$, with Y a complete uniform space, is prenear-standard in $CB(X, Y)$ relative to the strong uniformity if and only if f is microstep on X and the image of f consists of nearstandard points of Y .*

$\triangleleft \rightarrow$: Using 4.5.14 and 4.5.6, conclude that f is nearstandard in the strong uniformity. Therefore, there is some $g \in {}^\circ CB(X, Y)$ such that $f(x) \approx g(x)$ for all $x \in X$. Obviously, $\text{im } f \subset \approx \text{im}(g)$. Moreover, $\text{im}(g) \subset \text{pst}(Y)$ (see 4.5.13). If \mathcal{E} is a tiny cover then, by the definition of total boundedness, to each standard $V \in \mathcal{U}_Y$ there is a standard finite cover \mathcal{E}' of X satisfying $g(E)^2 \subset V$ for all $E \in \mathcal{E}'$. Therefore, $(\forall E \in \mathcal{E}) g(E)^2 \subset V$, i.e., g is nearstep on \mathcal{E} . Hence, $g(x) \approx f(x) \approx f(x') \approx g(x')$ for all $E \in \mathcal{E}$ and $x, x' \in E$, i.e., f is also nearstep relative to \mathcal{E} . Since \mathcal{E} is arbitrary, f is a microstep mapping.

\leftarrow : Since $\text{im}(f) \subset \text{nst}(Y)$; therefore,

$$(\forall x \in X)(\exists^{\text{st}} y \in Y)(\forall^{\text{st}} W \in \tau(y))(f(x) \in W).$$

By the construction principle 3.3.12,

$$(\forall^{\text{st}} W(\cdot))(\forall x \in X)(\exists^{\text{st}} y \in Y)(f(x) \in W(y)).$$

By idealization,

$$(\forall^{\text{st}} W(\cdot))(\exists^{\text{st}} \{y_1, \dots, y_n\})(\forall x \in X)(\exists k)(f(x) \in W(y_k)).$$

We now take $V \in \mathcal{U}_Y$. By hypothesis, for every tiny cover \mathcal{E} of X and for all $E \in \mathcal{E}$ we have $f(E)^2 \subset V$. Using the Cauchy principle 4.1.17 and recalling that the tiny covers are exactly the remote elements of the directed set of finite covers, we see that there is a standard finite cover \mathcal{E}_V satisfying $f(E)^2 \subset V$ for $E \in \mathcal{E}_V$.

Choose a corresponding standard cover \mathcal{E}_V and a standard finite set Y_0 of Y elements, for which $\text{im}(f) \subset V(Y_0)$.

Using \mathcal{E}_V and Y_0 , it is easy to construct a standard step function f_V such that $(\forall x \in X)((f_V(x), f(x)) \in V)$. Obviously, if $U \in {}^\circ\mathcal{U}_Y$ obeys the conditions $U = U^{-1}$ and $U \circ U \subset V$ then $(f_{V'}(x), f_{V''}(x)) \in V' \circ V''^{-1} \subset U \circ U \subset V$ for all $V', V'' \subset U$. Considering $\{f_V : V \in {}^\circ\mathcal{U}_Y\}$, note that $(f_V)_{V \in {}^\circ\mathcal{U}_Y}$ is a Cauchy net. Denote by g the standard limit of $(f_V)_{V \in {}^\circ\mathcal{U}_Y}$ in $CB(X, Y)$. We still see that $(\forall^{\text{st}} V \in \mathcal{U}_Y)(\forall x \in X)((g(x), f(x)) \in V)$. Consequently, $g \approx f$ in the strong uniformity. Therefore, f is nearstandard; hence, prenearstandard by completeness of $CB(X, Y)$ (cf. 4.5.14). \triangleright

4.5.17. Theorem. *If E is a subset of a complete separated space X then the following are equivalent:*

- (1) E is relatively compact;
- (2) E is precompact (i.e., the completion of E is compact);
- (3) E is totally bounded;
- (4) $E \subset \text{pst}(X)$;
- (5) $E \subset \text{nst}(X)$;
- (6) E lies in the microhalo of a finite set;
- (7) $\text{cl}(U)$ has a finite skeleton.

\triangleleft Since X is complete; by 4.5.6, $\text{pst}(X) = \text{nst}(X)$. Therefore, (5) \rightarrow (1) \rightarrow (4) (see 4.3.8). Obviously, (7) \rightarrow (6) \rightarrow (3) \rightarrow (1) \rightarrow (2). If (2) holds then $\text{cl}(E)$ is complete and totally bounded by the Hausdorff test. From 4.5.11 we infer the implication (2) \rightarrow (7). \triangleright

4.5.18. Theorem. *Assume that X is a compact space, Y is a complete uniform space, and $C(X, Y)$ is the space of continuous functions from X to Y endowed with the strong uniformity. For an internal member f of $C(X, Y)$ the following are equivalent:*

- (1) f is prenearstandard;
- (2) f is nearstandard;
- (3) f is microcontinuous and sends each standard point to a nearstandard point.

\triangleleft (1) \rightarrow (2): Obviously, f is prenearstandard in Y^X in the strong uniformity by 4.5.4, while f is nearstandard in Y^X by 4.5.8 and 4.5.6, i.e., there is a standard $g \in Y^X$ such that $f(x) \approx g(x)$ for all $x \in X$. Let $(f_\xi)_{\xi \in \Xi}$ be a standard net in $C(X, Y)$ microconvergent to f . Take $x' \approx x$ and note that $f_\xi(x') \approx f_\xi(x)$ for all standard $\xi \in \Xi$ (since f_ξ is continuous and X is compact). Then (cf. 3.3.17(3)) $f_\eta(x') \approx f_\eta(x)$ for some $\eta \in {}^a\Xi$. Hence, $g(x') \approx f(x') \approx f_\eta(x') \approx f_\eta(x) \approx f(x) \approx g(x)$. Therefore, the standard function g is microcontinuous and so $g \in CB(X, Y)$ by 4.4.6.

(2) \rightarrow (3): By hypothesis, there is a standard continuous function g such that $g(x) \approx f(x)$ for all $x \in X$. Therefore, $f(\circ X) \subset \approx g(\circ X) \subset \approx^\circ g(X) \subset \text{nst}(Y)$. Moreover, by 4.5.6, g is microcontinuous and so $f(x) \approx g(x) \approx g(x') \approx f(x')$ for all $x' \approx x$.

(3) \rightarrow (1): By 4.5.3, we are to demonstrate only that (3) \rightarrow (2). Choose a microcontinuous f such that $f(\circ X) \subset \text{nst}(X)$. By the construction principle 3.3.12, there is a standard function g such that $g(x) \in {}^\circ f(x)$ for all $x \in \circ X$. Check that g is uniformly continuous. To this end, choose a standard entourage $V \in \mathcal{U}_Y$ and a standard set $W \in \mathcal{U}_Y$ from the condition $W \circ W \circ W \subset V$. Using 4.5.7, find a standard U of the unique uniformity \mathcal{U}_X (see 4.4.8(4)) so that $f^\times(U) \subset W$. Considering standard points $x, x' \in \circ X$ satisfying $(x, x') \in U$, note that $(f(x), f(x')) \in W$, $(f(x'), g(x')) \in W$, and $(g(x), f(x)) \in W$. Therefore, $(g(x), g(x')) \in W \circ W \circ W \subset V$. Finally,

$$(\forall^{\text{st}} V \in \mathcal{U}_Y)(\exists^{\text{st}} U \in \mathcal{U}_X)(\forall^{\text{st}} x, x' \in U)((g(x), g(x')) \in V).$$

By transfer, we infer $g \in C(X, Y)$.

Given an arbitrary $x \in X$, observe now that $f(x) = f(x') \approx g(x') \approx g(x)$, where x' is the only standard point infinitely close to x . Hence, f is infinitely close to g in the strong uniformity. \triangleright

4.5.19. Ascoli–Arzelà Theorem. *Assume that X is a compact space, Y is a complete separated uniform space, and $E \subset C(X, Y)$. The set E is relatively compact in the strong uniformity if and only if E is equicontinuous and uniformly (totally) bounded (i.e., there is some totally bounded C in Y such that $f(X) \subset C$ for all $f \in E$).*

\triangleleft The claim follows from 4.5.18, 4.5.17, and 4.4.6 (2). \triangleright

4.6. Relative Monads

The notion of relatively standard element we have introduced in 3.9 is convenient for characterizing various topological properties.

4.6.1. Assume that τ is an arbitrary admissible element (see 3.9.2) and X is a τ -standard topological space. Given a τ -standard point $a \in X$, define the τ -monad of a as the intersection of τ -standard neighborhoods of a :

$$\mu^\tau(a) := \bigcap \{u \subset X : u \text{ is open; } a \in u; u \text{ st } \tau\}.$$

If $x \in \mu^\tau(a)$ then x is τ -infinitely close to a ; in symbols, $x \overset{\tau}{\approx} a$.

If X is a uniform space with τ -standard uniformity \mathcal{U} then x and y are τ -infinitely close points in X , in symbols, $x \overset{\tau}{\approx} y$ provided that $(x, y) \in \bigcap \{U \in \mathcal{U} : U \text{ st } \tau\}$.

If a is not an isolated point then $\mu^\tau(a) - \{a\} \neq \emptyset$ by idealization.

4.6.2. Assume that τ and λ are admissible elements; moreover, τ is λ -standard. If a topological space X and a point $a \in X$ are τ -standard and $\tau \text{ st } \lambda$, then

- (1) $\mu^\lambda(a) \subset \mu^\tau(a)$;
- (2) If $x \overset{\lambda}{\approx} y$ then $x \overset{\tau}{\approx} y$ for all $x, y \in X$;
- (3) If X and a are standard then $\mu^\lambda(a) \subset \mu(a)$;
- (4) If X and a are standard then $x \overset{\lambda}{\approx} y$ implies $x \approx y$ for all $x, y \in X$.

◁ The claim follows from 3.9.4 (2) also implying that $X \text{ st } \lambda$, i.e., the definition of $\mu^\lambda(a)$ is sound.

In case X and a are standard it suffices to put $\tau = \emptyset$ (or another standard set) instead of \emptyset since $\text{st}(X)$ and $X \text{ st } \emptyset$ are equivalent. ▷

4.6.3. Theorem. Assume given a topological space X , a subset A of X , and a point $a \in X$. Then

- (1) A is open if and only if $\mu^\tau(x) \subset A$ for all τ -standard $x \in A$;
- (2) A is closed if and only if each τ -standard point $x \in X$ τ -infinitely close to some point in A belongs to A : in symbols,

$$(\forall^{\text{st } \tau} x \in X)(\forall \eta \in A)(\eta \in \mu^\tau(x) \rightarrow x \in A).$$

◁ Everything is proven in much the same way as in the case of the standard environment. For completeness, we will demonstrate (1).

Let σ be the topology on X , and let $\sigma(a)$ be the collection of open neighborhoods of a . It follows from the τ -standardness of X that σ is τ -standard as well as $\sigma(a)$ is τ -standard for every τ -standard $a \in X$ by 3.9.7 (1).

Assume now that A is open, and take τ -standard element $x \in A$. By definition, $\mu^\tau(x) = \bigcap \{u : u \text{ st } \tau, u \in \sigma(x)\}$. Hence, $\mu^\tau(x) \subset A$ since $A \in \sigma(x)$.

To prove the converse, assume by way of contradiction that $\mu^\tau(x) \subset A$ for all τ -standard $x \in A$, but A is not open. Since X and A are τ -standard, by relative transfer principle we have

$$(\exists^{\text{st } \tau} x \in A)(\forall^{\text{st } \tau} U \in \sigma(x))(\exists^{\text{st } \tau} y)(y \in U \wedge y \notin A).$$

Consider the binary relation $\mathcal{R} \subset \sigma(x) \times X$ with

$$(u, z) \in \mathcal{R} \leftrightarrow z \in U \wedge z \notin A.$$

Since $\bigcap I \in \sigma(x)$ and $\bigcap I \text{ st } \tau$ for every τ -standard finite set $I \subset \sigma(x)$, the relation \mathcal{R} satisfies the condition of the idealization principle. Hence, $(\exists y)(\forall^{\text{st } \tau} U \in \sigma(x))(y \in U \wedge y \notin A)$ implying that $y \in \mu^\tau(x) - A$, which is a contradiction. ▷

4.6.4. Theorem. Assume that topological spaces X and Y , a mapping $f : X \rightarrow Y$ between them, and points $a \in X$ and $b \in Y$ are all τ -standard. If τ is a λ -standard element for some admissible λ then

- (1) $\lim_{x \rightarrow a} f(x) = b$ if and only if $\xi \overset{\lambda}{\approx} a$ implies $f(\xi) \overset{\tau}{\approx} b$ for all $\xi \in X$;
- (2) In case X and Y are uniform spaces, f is uniformly continuous if and only if $\xi \overset{\lambda}{\approx} \eta$ implies $f(\xi) \overset{\tau}{\approx} f(\eta)$ for all $\xi, \eta \in X$.

◁ We will prove (1). Using the τ -standardness of f , a , and b , infer on appealing to the relative transfer principle that

$$\lim_{x \rightarrow a} f(x) = b \leftrightarrow (\forall^{\text{st}} \tau W \in \sigma_Y(b))(\exists^{\text{st}} \tau u \in \sigma_X(a))(f(u) \subset W),$$

with σ_Y and σ_X the topologies of Y and X respectively. Agree that $\lim_{x \rightarrow a} f(x) = b$ and demonstrate that $f(\mu^\lambda(a)) \subset \mu^\tau(b)$. By 4.6.2 (1) $\mu^\lambda(a) \subset \mu^\tau(a)$, and so it suffices to show that $f(\mu^\tau(a)) \subset \mu^\tau(b)$ which amounts to

$$(\forall^{\text{st}} \tau W \in \sigma_Y(b))(f(\mu^\tau(a)) \subset W).$$

Suppose that $u \text{ st } \tau$, $u \in \sigma_X(a)$, and $f(u) \subset W$. Then $\mu^\tau(a) \subset u$ and so $f(\mu^\tau(a)) \subset W$, which completes the proof.

To demonstrate the converse, assume that $f(\mu^\lambda(a)) \subset \mu^\tau(b)$. Fix an arbitrary τ -standard neighborhood $W \in \sigma_Y(b)$ and note that $W \text{ st } \lambda$ and $f(\mu^\lambda(a)) \subset W$ by 3.9.4 (2).

Show first that $(\exists^{\text{st}} \lambda U \in \sigma_X(a))(f(U) \subset W)$. Were this false, the relation $\mathcal{R}_1 \subset \sigma_X(a) \times Y$ with $\mathcal{R}_1 := \{(U, y) : y \in U \wedge f(y) \notin W\}$ would satisfy the hypotheses of the relative idealization principle, implying

$$(\exists y)(\forall^{\text{st}} \lambda U \in \sigma_X(a))(y \in U \wedge f(y) \notin W).$$

This would contradict the inclusion $f(\mu^\lambda(a)) \subset W$.

Thus, $(\exists u \in \sigma_X(a))(f(u) \subset W)$. Since all parameters here are τ -standard, the relative transfer principle implies that $(\exists^{\text{st}} \tau U \in \sigma_X(a))(f(U) \subset W)$, which is what was required.

Item (2) is proved along the similar lines. ▷

4.6.5. Theorems 4.6.3 and 4.6.4 are applicable in the case of arbitrary admissible objects because we have $x \text{ st } x$ for every admissible x . For example, if $\tau := (X, A)$ (or if $\tau := (X, Y, f, a, b)$) then X and A in Theorem 4.6.3 (and X, Y, f, a , and b in Theorem 4.6.4) are τ -standard. In particular, this implies the following propositions.

- (1) Assume that X is an admissible topological space and $A \subset X$. If $\tau := (X, A)$ then A is open if and only if $\mu^\tau(x) \subset A$ for all τ -standard x in A .

- (2) Assume that X and Y are admissible topological spaces, $f: X \rightarrow Y$, $a \in X$, and $b \in Y$. If $\tau := (X, Y, f, a, b)$ then

$$\lim_{x \rightarrow a} f(x) = b \leftrightarrow (\forall x \in \mu^\tau(a)) f(x) \in \mu^\tau(b).$$

4.6.6. We now consider the case of $X := \mathbb{R}$ in more detail. Fix an internal admissible set τ . Let $x \in \mathbb{R}$ be an arbitrary (possibly nonstandard) real. We say that x is τ -infinitesimal or τ -infinitely small, in symbols: $x \overset{\tau}{\approx} 0$, provided that $(\forall^{\text{st}} \tau y \in \mathbb{R}_+) |x| < y$. As usual, the word “infinitesimal” serves as an adjective as well as a noun.

By analogy, we introduce the following natural definitions: x is a τ -unlimited or τ -infinitely large real, in symbols: $x \overset{\tau}{\approx} \infty$ or $x \overset{\tau}{\sim} \infty$, provided that $1/x$ is τ -infinitesimal; x is a τ -limited or τ -finite real provided that x is not τ -infinitely large, in symbols: $x \overset{\tau}{\ll} \infty$.

These definitions readily imply the following:

$$\begin{aligned} x \overset{\tau}{\approx} 0 &\leftrightarrow (\forall^{\text{st}} \tau y \in \mathbb{R}_+) (|x| \leq y), \\ x \overset{\tau}{\approx} \infty &\leftrightarrow (\forall^{\text{st}} \tau y \in \mathbb{R}_+) (|x| > y), \\ x \overset{\tau}{\ll} \infty &\leftrightarrow (\exists^{\text{st}} \tau y \in \mathbb{R}_+) (|x| \leq y). \end{aligned}$$

4.6.7. A real $x \in \mathbb{R}_+$ is τ -infinitesimal if and only if $|x| < \varphi(\tau)$ for each \mathbb{R}_+ -valued standard function φ satisfying $\tau \in \text{dom}(\varphi)$.

◁ Necessity is obvious. To prove sufficiency suppose that $y \in \mathbb{R}_+$ and $y \text{ st } \tau$. Show that $|x| < y$. Since $y \text{ st } \tau$; therefore, there exists a standard function ψ satisfying $y \in \psi(\tau)$, $\tau \in \text{dom}(\psi)$, and $\text{rng } \psi \subset \mathcal{P}_{\text{fin}}(\mathbb{R}_+)$ (as usual, $\mathcal{P}_{\text{fin}}(A)$ stands for the set of finite subsets of A).

By 3.9.3 we may define the standard function $\varphi: \text{dom}(\psi) \rightarrow \mathbb{R}_+$ by putting $\varphi(\alpha) := \min \psi(\alpha)$. By hypotheses, $|x| < \varphi(\tau)$ while $y \geq \varphi(\tau)$ by the construction of φ . ▷

4.6.8. If n is a natural then $n \text{ st } x$ for all $x \in \mathbb{R}_+$ such that $n \leq x$.

◁ Denote by m the integral part $[x]$ of x . Note that $m \text{ st } x$ and $n \leq m$. Consider the set $\overline{m} = \{0, 1, \dots, m\}$. Clearly, $\overline{m} \text{ st } m$ (Proposition 3.9.7(1)). Moreover, \overline{m} is a finite set, in symbols: $\text{fin}(\overline{m})$. By 3.9.7(2), $n \text{ st } \overline{m}$ whenever $n \in \overline{m}$. ▷

4.6.9. If $\lambda \text{ st } \tau$ and $x \overset{\tau}{\approx} 0$ ($x \overset{\tau}{\approx} \infty$) then $x \overset{\lambda}{\approx} 0$ ($x \overset{\lambda}{\approx} \infty$). If x is λ -limited then x is τ -limited.

◁ Immediate from 4.6.6. ▷

4.6.10. Theorem. *The following hold:*

- (1) **Limitedness Principle.** *If an internal set $B \subset \mathbb{R}$ consists only of τ -limited members then there is a τ -standard $t \in \mathbb{R}$ satisfying $B \subset [-t, t]$.*
- (2) **Permanence Principle.** *If an internal set B contains all positive τ -limited reals then B includes the interval $[0, \Omega]$ with Ω some τ -unlimited real Ω .*
- (3) **Cauchy Principle.** *If an internal set B contains all τ -infinitesimals then B includes the interval $[-a, a]$ with a some τ -standard positive real.*
- (4) **Robinson Principle.** *If an internal set B consists only of τ -infinitesimals then B is included in the interval $[-\varepsilon, \varepsilon]$ with ε some positive τ -infinitesimal.*

◁ We confine demonstration to items (1) and (4) since (2) and (3) are proved similarly.

(1): If $\lambda \approx^{\tau} \infty$ then $|\xi| < \lambda$ for all $\xi \in B$ by hypothesis (cf. 4.6.6). In other words, B is a bounded set. The set $B' := \{|b| : b \in B\}$ is bounded above as so B' has the positive least upper bound $\mu := \sup(B')$. If $\mu \approx^{\tau} \infty$ then $\mu - 1 \approx^{\tau} \infty$, and $\mu - 1 < |\xi| \leq \mu$ for some $\xi \in B$ by definition. We arrive at a contradiction implying that ξ is limited. Consequently, μ is a τ -limited real, and so there is a τ -standard real $t \in \mathbb{R}$ such that $t > \mu$ (cf. 4.6.6). Obviously, $B \subset [-t, t]$.

(4): Take an arbitrary τ -standard real $y \in \mathbb{R}$. By hypothesis, $\xi \leq y$ for all $\xi \in B'$. Hence, B' is a bounded set. Put $\varepsilon := \sup(B')$. Note that $B \subset [-\varepsilon, \varepsilon]$ and $\varepsilon \leq y$ for all τ -standard $y \in \mathbb{R}$. ▷

4.6.11. *To each unlimited (infinitesimal) real $x \in \mathbb{R}$ there is a nonstandard real η such that $x \approx^{\eta} \infty$ ($x \approx^{\eta} 0$, respectively). This η may be chosen to be limited, or infinitesimal, or unlimited.*

◁ Consider the internal relation $\sigma \subset \mathbb{R}_+^{\mathbb{R}} \times \mathbb{R} \times \mathbb{R}$ with

$$(f, \xi, \eta) \in \sigma \leftrightarrow f(\eta) < |x| \wedge \eta \neq \xi.$$

If x is unlimited then it is easy to see that σ satisfies the hypotheses of the idealization principle:

$$(\forall^{\text{st fin}} M \subset \mathbb{R}_+^{\mathbb{R}} \times \mathbb{R})(\exists \eta)(\forall (f, \xi) \in M)((f, \xi, \eta) \in \sigma).$$

By idealization, we thus conclude that

$$(\exists \eta)(\forall^{\text{st}} f \in \mathbb{R}_+^{\mathbb{R}})(\forall^{\text{st}} \xi \in \mathbb{R})(f(\eta) < |x| \wedge \eta \neq \xi).$$

Such an η will suffice. If the so-obtained η is limited then $\eta' = \eta - \circ\eta \approx 0$ still meets the claims by 4.6.9. ▷

We now exhibit some examples that show how to obviate obstacles we have discussed in the beginning of 3.9.

4.6.12. Theorem. Suppose that $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ and $a \in \mathbb{R}$ are standard, and $\lim_{y \rightarrow 0} f(x, y)$ exists for each x in some neighborhood of zero. Then

$$\lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x, y) = a \leftrightarrow (\forall \alpha \approx 0)(\forall \beta \overset{\alpha}{\approx} 0)(f(\alpha, \beta) - a \approx 0).$$

◁ Put $a := \lim_{x \rightarrow 0} \lim_{y \rightarrow 0} f(x, y)$, and let

$$g(x) := \lim_{y \rightarrow 0} f(x, y).$$

Then $g(\alpha) \approx a$ for all $\alpha \approx 0$. Observe that g is a standard function by 3.9.3, and so $g(\alpha) \text{ st } \alpha$. By 4.6.2 and 3.9.4 (2), the equality $g(\alpha) = \lim_{y \rightarrow 0} f(\alpha, y)$ amounts to

$$(\forall \beta \overset{\alpha}{\approx} 0) f(\alpha, \beta) \overset{\alpha}{\approx} g(\alpha).$$

By 4.6.2, the approximate equality $f(\alpha, \beta) \approx g(\alpha)$ follows from $f(\alpha, \beta) \overset{\alpha}{\approx} \varphi(\alpha)$. Since $g(\alpha) \approx a$; therefore, $f(\alpha, \beta) \approx a$.

To prove the converse it suffices obviously to demonstrate that

$$(\forall \varepsilon > 0)(\exists \delta)(\forall x)(|x| < \delta \rightarrow (\exists \gamma)(\forall y)(|y| < \gamma \rightarrow |f(x, y) - a| < \varepsilon)).$$

To this end, take an arbitrary standard ε and consider the internal set

$$M := \{\delta > 0 : (\forall x)(|x| < \delta \rightarrow (\exists \gamma)(\forall y)(|y| < \gamma \rightarrow |f(x, y) - a| < \varepsilon))\}.$$

It is easy that M contains all infinitesimals. Indeed, if $\delta \approx 0$ and $|x| < \delta$ then $x \approx 0$. If $\gamma \overset{x}{\approx} 0$ then $(\forall y)(|y| < \gamma \rightarrow y \overset{x}{\approx} 0)$, so that $|f(x, y) - a| < \varepsilon$. By the Cauchy principle M obviously contains some standard element. ▷

This theorem holds also in the case when $x \rightarrow b$ and $y \rightarrow c$ for arbitrary standard b and c . It translates routinely to infinite limits and limits at infinity as well as to arbitrary topological spaces.

4.6.13. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is Riemann-integrable on each bounded interval and the integral $\int_{-\infty}^{\infty} f(x) dx$ exists at least in the principal value sense, then

$$\int_{-\infty}^{\infty} f(x) dx = \overset{\circ}{\left(\Delta \sum_{-\lceil \frac{a}{\Delta} \rceil}^{\lceil \frac{a}{\Delta} \rceil} * f(k\Delta) \right)}$$

for all $a \approx \infty$ and $\Delta \overset{a}{\approx} 0$.

◁ Immediate from 4.6.12. ▷

4.6.14. It stands to reason now to exhibit three simple examples.

(1) We first attempt at proving the “difficult” part of l’Hôpital’s rule.

Suppose that f and g are standard functions differentiable in a neighborhood of a standard point a . Suppose further that $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = \infty$, while $g'(x) \neq 0$ in a neighborhood of a and

$$\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} = d.$$

The task is to show that $\lim_{x \rightarrow a} f(x)/g(x) = d$. Take arbitrary points y and $z \approx a$. For definiteness, assume that $a < z < y$. By Cauchy’s Theorem, there is a point η in $[z, y]$ satisfying

$$\frac{{}^*f(y) - {}^*f(z)}{{}^*g(y) - {}^*g(z)} = \frac{{}^*f'(\eta)}{{}^*g'(\eta)} \approx d$$

because $\eta \approx a$.

We now consider the equality

$$\frac{{}^*f(y) - {}^*f(z)}{{}^*g(y) - {}^*g(z)} = \frac{{}^*f(z)}{{}^*g(z)} \left(1 - \frac{{}^*f(y)}{{}^*f(z)}\right) \left(1 - \frac{{}^*g(y)}{{}^*g(z)}\right)^{-1}.$$

This formula shows that if $z \overset{y}{\approx} a$ then ${}^*f(y)/{}^*f(z) \approx 0$ and ${}^*g(y)/{}^*g(z) \approx 0$ by Theorem 4.6.4 (1) (more precisely, by its obvious translation to infinite limits). Hence, ${}^*f(z)/{}^*g(z) \approx d$. Thus, ${}^*f(z)/{}^*g(z) \approx d$ for all $z \overset{y}{\approx} a$. By 4.6.4 (1), this means that $\lim_{x \rightarrow a} f(x)/g(x) = d$.

(2) We now complete the proof of the claim of 4.6.4 (2). To this end, take x' and x'' so that $x' \overset{N}{\approx} x''$. By 4.6.4 (2), ${}^*f_N(x') \approx {}^*f_N(x'')$ and we readily find that ${}^*f(x') \approx {}^*f(x'')$. Therefore, for all x' and x'' from $x' \overset{N}{\approx} x''$ it follows that ${}^*f(x') \approx {}^*f(x'')$. This amounts to the uniform continuity of f by 4.6.4 (2).

(3) As an example of application of 4.6.11 we consider the following proposition (cf. [3, 1.3.2]).

Let $(a_{m,n})_{m \in \mathbb{N}, n \in \mathbb{N}}$ be a standard double sequence having the limits

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} a_{m,n} = a; \quad \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} a_{m,n} = b.$$

Then

$$(\forall m \approx \infty)(\exists n_1, n_2 \approx \infty) \\ ((\forall n < n_1)(n \approx \infty \rightarrow a_{m,n} \approx a) \wedge (\forall n > n_2)(a_{m,n} \approx b)).$$

To prove, take $n_1 \approx \infty$ so that $m \overset{n_1}{\approx} \infty$ (this is possible by 4.6.11) and $n_2 \overset{m}{\approx} \infty$.

4.6.15. Theorem. *There are an unlimited natural N and a real x in $[0, 1]$ such that if y is N -infinitely close to x then y fails to be N -standard.*

◁ Proceed by contradiction and suppose that the claim is false. By 4.6.6, we then obtain the following theorem of IST

$$(\forall N \in \mathbb{N})(\forall x \in [0, 1])(\exists^{\text{st}} \varphi \in (\mathcal{P}_{\text{fin}}(\mathbb{R})_+)^{\mathbb{N}})(\exists z \in \mathbb{R}_+) \\ (\forall^{\text{st}} \psi \in \mathbb{R}_+^{\mathbb{N}})(z \in \varphi(N) \wedge |x - z| < \psi(N)).$$

Nelson's algorithm applies to the last formula. We agree to assume that the variables N, x, φ, z , and ψ range over the sets indicated in the formula by implication. Using the idealization principle, we arrive at the formula

$$(\forall N)(\forall x)(\exists^{\text{st}} \varphi)(\forall^{\text{st}} \Xi)(\exists z)(\forall \psi \in \Xi)(z \in \varphi(N) \wedge |x - z| < \psi(N)),$$

with $\Xi \in \mathcal{P}_{\text{fin}}(\mathbb{R}_+^{\mathbb{N}})$.

Applying the standardization principle, proceed to the equivalent formula

$$(\forall N)(\forall x)(\forall^{\text{st}} \tilde{\Xi})(\exists^{\text{st}} \varphi)(\exists z)(\forall \psi \in \tilde{\Xi}(\varphi))(z \in \varphi(N) \wedge |x - z| < \psi(N)),$$

where $\tilde{\Xi} : \mathcal{P}_{\text{fin}}(\mathbb{R}_+)^{\mathbb{N}} \rightarrow \mathcal{P}_{\text{fin}}(\mathbb{R}_+^{\mathbb{N}})$. Drag out the universal quantifiers and apply the idealization and transfer principles successively to come to the equivalent formula

$$(\forall \tilde{\Xi})(\exists \Phi)(\forall N)(\forall x)(\exists \varphi \in \Phi)(\forall \psi \in \tilde{\Xi}(\varphi))(\exists z) \\ (z \in \varphi(N) \wedge |x - z| < \psi(N)),$$

with $\Phi \in \mathcal{P}_{\text{fin}}(\mathcal{P}_{\text{fin}}(\mathbb{R}_+)^{\mathbb{N}})$. We are left with refuting the last formula in ZFC.

Define the function $\tilde{\Xi}$ and the set $M_{\varphi, n}$ as follows:

$$\tilde{\Xi}(\varphi) := \{\psi\}, \quad \psi(n) := a_n := 1/2n|\varphi(n)| \varphi \in \mathcal{P}_{\text{fin}}(\mathbb{R}_+)^{\mathbb{N}}; \\ M_{\varphi, n} := \bigcup_{z \in \varphi(n)} (z - a_n, z + a_n),$$

with $|\varphi(n)|$ the size of $\varphi(n)$. Note that $\nu(M_{\varphi, n}) \leq 1/n$, with ν standing for Lebesgue measure. In case $\Phi \in \mathcal{P}_{\text{fin}}(\mathcal{P}_{\text{fin}}(\mathbb{R}_+)^{\mathbb{N}})$ we then put

$$\overline{M}_{\Phi, n} := \bigcup_{\varphi \in \Phi} M_{\varphi, n}.$$

Clearly, $\nu(\overline{M}_{\Phi, n}) \leq |\Phi|/n$.

If the formula in question were true in ZFC then we would apply it to the above-constructed function to obtain

$$(\exists \Phi \in \mathcal{P}^{\text{fin}}(\mathcal{P}^{\text{fin}}(\mathbb{R}_+)^{\mathbb{N}}))(\forall n \in \mathbb{N})([0, 1] \subset \overline{M}_{\Phi, n}).$$

This is a contradiction since we would have $\nu(\overline{M}_{\Phi, n}) < 1$ for $n > |\Phi|$. ▷

4.6.16. COMMENTS.

(1) This section is an excerpt from [141] (cf. [146]).

(2) From 4.6.14 it follows also that the “infinitesimal” tests 4.3.6 do not admit abstraction to the case of τ -standard objects in contrast to the test 4.4.6 (1) (see 4.6.4).

(3) Consider the relation of strict standardness sst we have discussed in 3.9.16 (4). Replacing $\cdot \text{st} \cdot$ with $\cdot \text{sst} \cdot$ in 4.6.6, we define τ -infinitesimal (τ -unlimited or τ -limited) relative to the predicate $\cdot \text{sst} \cdot$. In this event, we see from 4.6.7 that the concept of τ -infinitesimal (τ -unlimited or τ -limited) real relative to $\cdot \text{st} \tau \cdot$ coincides with the respective concept relative to $\cdot \text{sst} \tau \cdot$.

(4) Slight modification of the proof of 4.6.7 shows that the concepts of τ -infinite proximity relative to $\cdot \text{st} \cdot$ and $\cdot \text{sst} \cdot$ coincide in arbitrary topological and uniform spaces. Therefore, Theorems 4.6.3, 4.6.4, 4.6.10, and 4.6.12 as well as Propositions 4.6.9 and 4.6.11 remain valid on substituting $\cdot \text{sst} \tau \cdot$ for $\cdot \text{st} \tau \cdot$.

(5) In contrast to (4), Proposition 4.6.8 fails for the predicate $\cdot \text{sst} \tau \cdot$. This entails that neither 3.9.4 (3) nor the implication \leftarrow in the relative idealization principle hold on replacing $\cdot \text{st} \tau \cdot$ with $\cdot \text{sst} \tau \cdot$. See [141, 146] for more details.

4.7. Compactness and Subcontinuity

This section collects standard and nonstandard tests for a filter to be compact and addresses related matters. These tests supplement the similar topological facts for sets as presented in 4.3 and 4.5. We give a few applications to the theory of subcontinuous correspondences which was suggested in [125, 450].

4.7.1. A filter \mathcal{F} (on a topological space X) is *compact* (cf. [392]) provided that each filter on X finer than \mathcal{F} has an adherent point in X . Similarly, a net is *compact* provided that its every subnet has a convergent subnet.

4.7.2. A standard filter \mathcal{F} on X is compact if and only if every member of the monad of \mathcal{F} is nearstandard: $\mu(\mathcal{F}) \subset \text{nst}(X)$.

$\triangleleft \rightarrow$: Take $x \in \mu(\mathcal{F})$. Consider the ultrafilter $(x) := * \{U \subset X : x \in U\}$ on the original space X . Clearly, $(x) \supset \mathcal{F}$ and so there is a standard point x' satisfying $x \approx x'$. In other words, x is a nearstandard point.

\leftarrow : If $\mathcal{G} \supset \mathcal{F}$ then $\mu(\mathcal{G}) \subset \mu(\mathcal{F})$. Take $x \in \mu(\mathcal{G})$. Then $x \in \text{nst}(X)$, i.e. $x \approx x'$ for some $x' \in {}^\circ X$, which means that x' is an adherent point of \mathcal{F} . \triangleright

4.7.3. A filter \mathcal{F} on X is compact if and only if to each open cover \mathcal{E} of X there is a finite subset of \mathcal{E} covering some member of \mathcal{F} .

$\triangleleft \rightarrow$: It suffices to work on assuming the standard environment. In this event if \mathcal{F} is compact then $\mu(\mathcal{F}) \subset \text{nst}(X)$. Since $\text{nst}(X)$ lies in the monad of every

standard cover \mathcal{E} of X ; therefore, $(\exists F \in \mathcal{F})(\forall x \in F)(\exists E \in {}^\circ\mathcal{E})(x \in E)$. Take as a sought F an arbitrary infinitesimal member of \mathcal{F} . Using the principles of transfer and idealization successively, complete the proof.

\leftarrow : Assume that \mathcal{E} is an open cover of X and $\mu(\mathcal{E})$ stands for the monad of \mathcal{E} ; i.e., by transfer there are a standard member F of $\in \mathcal{F}$ and standard finite subset \mathcal{E}_0 of \mathcal{E} satisfying $\bigcup \mathcal{E}_0 \supset F \supset \mu(\mathcal{F})$. Consequently, $\mu(\mathcal{F}) \subset \mu(\mathcal{E})$. We are done on recalling that $\text{nst}(X)$ is exactly the intersection of the monads of standard open covers of X . \triangleright

4.7.4. Proposition 4.7.3 makes it natural to seek for an analog for filters of the celebrated Hausdorff test for a set to be compact. To this end, we will work in a uniform space (X, \mathcal{U}) .

4.7.5. A filter \mathcal{F} on X is *totally bounded* provided that to each entourage $U \in \mathcal{U}$ there is a finite U -net for some member of \mathcal{F} .

4.7.6. A filter \mathcal{F} on X is *complete* provided that every Cauchy filter finer than \mathcal{F} converges in X .

4.7.7. A standard filter \mathcal{F} is *complete* if and only if every prenearstandard point of the monad of \mathcal{F} is nearstandard.

$\triangleleft \rightarrow$: Let $x \in \text{pst}(X) \cap \mu(\mathcal{F})$ be a prenearstandard point of \mathcal{F} . This means that x belongs to the monad of some Cauchy filter \mathcal{G} . In this event $\mu(\mathcal{F}) \cap \mu(\mathcal{G}) \neq \emptyset$. The least upper bound of \mathcal{G} and \mathcal{F} is clearly a Cauchy filter. Hence, there is a point x' in ${}^\circ X$ satisfying $x' \in \mu(\mathcal{G}) \cap \mu(\mathcal{F})$. Consequently, $x' \approx x$ and $x \in \text{nst}(X)$.

\leftarrow : Assume that $\mathcal{G} \supset \mathcal{F}$ and \mathcal{G} is a Cauchy filter. If $x \in \mu(\mathcal{G})$ then $x \in \mu(\mathcal{F}) \subset \text{nst}(X)$. This means that \mathcal{G} has an adherent point. \triangleright

4.7.8. A standard filter \mathcal{F} is *totally bounded* if and only if every point of the monad of \mathcal{F} is prenearstandard.

$\triangleleft \rightarrow$: By transfer, to each standard member U of the uniformity \mathcal{U} of the space X under study, there are a standard member F of \mathcal{F} and standard finite set E satisfying $U(E) \supset F$. Thus, $\mu(\mathcal{F}) \subset U(E)$ and so to each $x \in \mu(\mathcal{F})$ and each $U \in {}^\circ\mathcal{U}$ there is some standard x' such that $x \in U(x')$. Put $\mathcal{G} := \{U(x') : U \in \mathcal{U}, x \in U(x')\}$. Clearly, \mathcal{G} is a base for a Cauchy filter and $x \in \mu(\mathcal{G})$ by construction. Consequently, $\mu(\mathcal{F}) \subset \text{pst}(X)$.

\leftarrow : Assume by way of contradiction that $\mu(\mathcal{F}) \subset \text{pst}(X)$, but \mathcal{F} is not totally bounded. By transfer, there is a standard member U of ${}^\circ\mathcal{U}$ such that to each $F \in {}^\circ\mathcal{F}$ and each standard finite E there is some x in F in the complement of $U(E)$. By idealization, there is a point x in $\mu(\mathcal{F})$ such that $x \notin U(y)$ for all standard $y \in X$. By hypothesis $x \in \mu(\mathcal{G})$, with \mathcal{G} a Cauchy filter. Take $G \in {}^\circ\mathcal{G}$ so that $G \times G \subset U$. Note that $x \in \mu(\mathcal{G}) \subset U(y)$ for all $y \in G$, which is a contradiction. \triangleright

4.7.9. Theorem. *A filter is compact if and only if it is complete and totally bounded.*

$\triangleleft \rightarrow$: It suffices to assume the standard environment. If \mathcal{F} is a compact filter then $\mu(\mathcal{F}) \subset \text{nst}(X)$ by 4.7.2. Since $\text{nst}(X) \subset \text{pst}(X)$, conclude that \mathcal{F} is complete and totally bounded.

\leftarrow : Since \mathcal{F} is a totally bounded filter; therefore, $\mu(\mathcal{F}) \subset \text{pst}(X)$ by 4.7.8. Since \mathcal{F} is a complete filter, $\mu(\mathcal{F}) \cap \text{pst}(X) \subset \text{nst}(X)$. Consequently, $\mu(\mathcal{F}) = \mu(\mathcal{F}) \cap \text{pst}(X) \subset \text{nst}(X)$. The proof is complete on recalling 4.7.2. \triangleright

4.7.10. The above is helpful in studying various topological concepts that are close to continuity. We will address one of these concepts (cf. [125, 392, 450]).

4.7.11. A correspondence Γ from X to Y is *subcontinuous at a point* x of $\text{dom}(\Gamma)$ provided that the image of the neighborhood filter of x under Γ is a compact filter on Y . A correspondence Γ is *subcontinuous* provided that Γ is subcontinuous at every point of $\text{dom}(\Gamma)$.

4.7.12. *A standard correspondence Γ from X to Y is subcontinuous if and only if $\Gamma(\text{nst}(X)) \subset \text{nst}(Y)$.*

\triangleleft The claim follows from 4.7.2, since $\text{nst}(X)$ is the union of monads of points in the standard part ${}^\circ X$ of X . \triangleright

4.7.13. *A correspondence is subcontinuous if and only if it sends each compact filter to a compact filter.*

\triangleleft Since the neighborhood filter of every point is always compact, the necessity part is beyond a doubt. Assume now that we deal with a subcontinuous correspondence. Without loss of generality, we will proceed in the standard environment. Using 4.7.12 and 4.7.2, note that the image of a standard compact filter is compact too. It suffices to appeal to the transfer principle. \triangleright

4.7.14. In view of 4.7.13 a subcontinuous correspondence is sometimes referred to as a *compact correspondence* (cf. [392]).

4.7.15. *Every subcontinuous correspondence with range a Hausdorff space preserves relative compactness.*

\triangleleft If U is a standard relatively compact set in X then $U \subset \text{nst}(X)$. Hence, $\Gamma(U) \subset \text{nst}(Y)$. By 4.3.8 $\Gamma(U)$ is relatively compact. \triangleright

4.7.16. *Let Γ be a closed subcontinuous mapping. Then Γ is upper semicontinuous.*

\triangleleft By transfer, assume the standard environment. Let A be a standard closed set and $x \in \text{cl}(\Gamma^{-1}(A))$. There is a point x' infinitely close to x such that $(x', a') \in \Gamma$ for some $a' \in A$.

Since $a' \in \Gamma(\text{nst}(X))$, there is a standard point a in the target of Γ such that $a \approx a'$. Since A is closed; therefore, $a \in A$. Since Γ is closed; therefore, $(x, a) \in \Gamma$. All in all, $x \in \Gamma^{-1}(A)$. \triangleright

4.7.17. Proposition 4.7.16 is established in [450] and generalizes an earlier fact about functions in [125]. In closing, we give an easy proof for a slight modification of the test 5.1 of [125].

4.7.18. *Let $f : X \rightarrow Y$ be a function with range a Hausdorff space. Then f is continuous if and only if to each point x in X there is an element y in Y such that the condition $x_\xi \rightarrow x$ implies existence of a subnet $(x_\eta)_{\eta \in \mathbb{H}}$ satisfying $f(x_\eta) \rightarrow y$.*

\triangleleft Only the sufficiency part needs proving. Assume the standard environment. By hypothesis

$$(\forall x_\xi \rightarrow x)(\exists y_\eta \rightarrow y)(x_\eta, y_\eta) \in f,$$

which may easily be rewritten as

$$(\forall x' \approx x)(\exists y' \approx y)(x', y') \in f.$$

In particular, $y' = f(x)$ for some $y' \approx y$. Since Y is Hausdorff, $y = f(x)$. Moreover, $x' \approx x \rightarrow f(x') \approx f(x)$, i.e. f is a continuous function. \triangleright

4.8. Cyclic and Extensional Filters

This section contains (mostly easy) prerequisites for descending and ascending filters.

4.8.1. *If $(A_\xi)_{\xi \in \Xi}$ is a family of nonempty members of $\mathbb{V}^{(B)}$ and $(b_\xi)_{\xi \in \Xi}$ is a partition of unity then*

$$\left(\sum_{\xi \in \Xi} b_\xi A_\xi \right) \downarrow = \sum_{\xi \in \Xi} b_\xi A_\xi \downarrow.$$

\triangleleft Put $A := \sum_{\xi \in \Xi} b_\xi A_\xi$. Given $\xi \in \Xi$, observe immediately that $\llbracket a \in A_\xi \rrbracket \geq \llbracket a \in A \rrbracket \wedge \llbracket A = A_\xi \rrbracket = \llbracket A = A_\xi \rrbracket \geq b_\xi$, whenever $a \in A \downarrow$. Using the transfer principle for $\mathbb{V}^{(B)}$, derive $\llbracket a \in A_\xi \rrbracket = \llbracket (\exists a_\xi \in A_\xi)(a = a_\xi) \rrbracket$. Appealing to the maximum principle, infer $(\exists a_\xi \in A_\xi \downarrow) \llbracket a \in A_\xi \rrbracket = \llbracket a = a_\xi \rrbracket \geq b_\xi$. Therefore, $a = \sum_{\xi \in \Xi} b_\xi a_\xi$.

Assume now that $b_\xi a = b_\xi a_\xi$ for some $a_\xi \in A_\xi \downarrow$ and all $\xi \in \Xi$. Since $\llbracket A = A_\xi \rrbracket \geq b_\xi$ ($\xi \in \Xi$) by the definition of mixing, we then derive that $\llbracket a \in A \rrbracket \geq \llbracket a = a_\xi \rrbracket \wedge \llbracket a_\xi \in A_\xi \rrbracket \wedge \llbracket A_\xi = A \rrbracket \geq b_\xi$ for $\xi \in \Xi$, i.e., $\llbracket a \in A \rrbracket \geq \bigvee_{\xi \in \Xi} b_\xi = 1$ and $a \in A \downarrow$. \triangleright

4.8.2. If $A_\xi \in \mathcal{P}(\mathbb{V}^{(B)})$ is a cyclic set for every $\xi \in \Xi$ then

$$\sum_{\xi \in \Xi} b_\xi A_\xi \uparrow = \left(\sum_{\xi \in \Xi} b_\xi A_\xi \right) \uparrow.$$

◁ Since $A_\xi \uparrow \downarrow = A_\xi$ for $\xi \in \Xi$ by hypothesis, from 4.8.1 we infer

$$\left(\sum_{\xi \in \Xi} b_\xi A_\xi \uparrow \right) \downarrow = \sum_{\xi \in \Xi} b_\xi A_\xi \uparrow \downarrow = \sum_{\xi \in \Xi} b_\xi A_\xi.$$

Recall that if A is a nonempty set inside $\mathbb{V}^{(B)}$ then $A = A \downarrow \uparrow$. Consequently,

$$\sum_{\xi \in \Xi} b_\xi A_\xi \uparrow = \left(\sum_{\xi \in \Xi} b_\xi A_\xi \uparrow \right) \downarrow \uparrow = \left(\sum_{\xi \in \Xi} b_\xi A_\xi \right) \uparrow,$$

which completes the proof. ▷

4.8.3. Let $(b_\xi)_{\xi \in \Xi}$ be a partition of unity. Assume further that $(X_\xi)_{\xi \in \Xi}$ and $(Y_\xi)_{\xi \in \Xi}$ are some families satisfying $\llbracket X_\xi \supset Y_\xi \rrbracket = \mathbb{1}$ for all $\xi \in \Xi$. Then

$$\left[\left[\sum_{\xi \in \Xi} b_\xi X_\xi \supset \sum_{\xi \in \Xi} b_\xi Y_\xi \right] \right] = \mathbb{1}.$$

◁ Put $X := \sum_{\xi \in \Xi} b_\xi X_\xi$ and $Y := \sum_{\xi \in \Xi} b_\xi Y_\xi$. Clearly, $\llbracket Y \subset X \rrbracket \geq \llbracket X = X_\xi \rrbracket \wedge \llbracket X_\xi \supset Y \rrbracket \geq \llbracket X = X_\xi \rrbracket \wedge \llbracket X_\xi \supset Y_\xi \rrbracket \wedge \llbracket Y = Y_\xi \rrbracket \geq b_\xi \wedge \mathbb{1} \wedge b_\xi = b_\xi$ for all $\xi \in \Xi$. ▷

4.8.4. Let X be a nonempty member of $\mathbb{V}^{(B)}$. Then

$$\llbracket \mathcal{P}_{\text{fin}}(X) = \mathcal{P}_{\text{fin}}(X \downarrow)^\uparrow \rrbracket = \mathbb{1},$$

where $\mathcal{P}_{\text{fin}}(A)$ stands as usual for the collection of finite subsets of A and $\mathcal{P}_{\text{fin}}(X \downarrow)^\uparrow := \{Y \uparrow : Y \in \mathcal{P}_{\text{fin}}(X \downarrow)\} \uparrow$.

◁ The inclusion $\mathcal{P}_{\text{fin}}(X \downarrow)^\uparrow \subset \mathcal{P}_{\text{fin}}(X)$ raises no doubt inside $\mathbb{V}^{(B)}$ (since the descent of a finite set is finite too). We are thus left with checking the relation

$$\llbracket (\forall t) t \in \mathcal{P}_{\text{fin}}(X) \rightarrow t \in \mathcal{P}_{\text{fin}}(X \downarrow)^\uparrow \rrbracket = \mathbb{1}.$$

This amounts to the equality

$$\bigwedge \{ \llbracket t \in \mathcal{P}_{\text{fin}}(X \downarrow)^\uparrow \rrbracket : \llbracket t \in \mathcal{P}_{\text{fin}}(X) \rrbracket = \mathbb{1} \} = \mathbb{1}.$$

If $\llbracket t \in \mathcal{P}_{\text{fin}}(X) \rrbracket = \mathbb{1}$ then the transfer principle yields

$$\mathbb{1} = \llbracket (\exists n \in \mathbb{N}^\wedge)(\exists f : n \rightarrow \mathbb{X})(t = \text{im}(f)) \rrbracket = \bigvee_{n \in \mathbb{N}} \llbracket f : n^\wedge \rightarrow \mathbb{X} \rrbracket \wedge \llbracket t = \text{im}(f) \rrbracket.$$

Using the mixing and maximum principles, find a countable partition of unity (b_n) in B and a sequence (f_n) in $\mathbb{V}^{(B)}$ so that $b_n \leq \llbracket f_n : n^\wedge \rightarrow X \rrbracket \wedge \llbracket t = \text{im}(f) \rrbracket$. Without loss of generality, we may assume that $\llbracket f_n : n^\wedge \rightarrow X \rrbracket = \mathbb{1}$. Put $g_n := f \downarrow : n \rightarrow X \downarrow$. Then $\text{im}(g) \in \mathcal{P}_{\text{fin}}(X \downarrow)$ and $b_n \leq \llbracket t = (\text{im}(g)) \uparrow \rrbracket$. Therefore,

$$\begin{aligned} \mathbb{1} &= \bigvee_{n \in \mathbb{N}} \llbracket t = (\text{im}(g)) \uparrow \rrbracket \leq \bigvee \{ \llbracket t = u \rrbracket : u \in \mathcal{P}_{\text{fin}}(X \downarrow) \uparrow \} \\ &= \llbracket t \in \mathcal{P}_{\text{fin}}(X \downarrow) \uparrow \uparrow \rrbracket, \end{aligned}$$

which completes the proof. \triangleright

4.8.5. Let \mathcal{G} be a filterbase on X , with X a subset of $\mathbb{V}^{(B)}$, i.e., $X \in \mathcal{P}(\mathbb{V}^{(B)})$. Put

$$\begin{aligned} \mathcal{G}' &:= \{ F \in \mathcal{P}(X \uparrow) \downarrow : (\exists G \in \mathcal{G}) \llbracket F \supset G \uparrow \rrbracket = \mathbb{1} \}; \\ \mathcal{G}'' &:= \{ G \uparrow : G \in \mathcal{G} \}. \end{aligned}$$

Then $\mathcal{G}' \uparrow$ and $\mathcal{G}'' \uparrow$ are bases for the same filter $\mathcal{G}' \uparrow$ on $X \uparrow$ inside $\mathbb{V}^{(B)}$.

\triangleleft Check that $\mathcal{G}' \uparrow$ is a filterbase on $X \uparrow$ inside $\mathbb{V}^{(B)}$. To this end note that

$$\begin{aligned} &\llbracket (\forall F_1, F_2 \in \mathcal{G}' \uparrow)(\exists F \in \mathcal{G}' \uparrow)(F \subset F_1 \cap F_2) \rrbracket \\ &= \bigwedge_{F_1, F_2 \in \mathcal{G}' \uparrow} \llbracket (\exists F \in \mathcal{G}' \uparrow)(F \subset F_1 \subset F_2) \rrbracket. \end{aligned}$$

If $F_1, F_2 \in \mathcal{G}' \uparrow$ then there are $G_1, G_2 \in \mathcal{G}$ such that $\llbracket F_1 \supset G_1 \uparrow \rrbracket = \mathbb{1}$ and $\llbracket F_2 \supset G_2 \uparrow \rrbracket = \mathbb{1}$. Take $G \in \mathcal{G}$ such that $G \subset G_1 \cap G_2$ to find $(G_1 \cap G_2) \uparrow \in \mathcal{G}' \uparrow$ and

$$\llbracket F_1 \cap F_2 \supset (G_1 \cap G_2) \uparrow \rrbracket \geq \llbracket F_1 \supset G_1 \uparrow \rrbracket \wedge \llbracket F_2 \supset G_2 \uparrow \rrbracket = \mathbb{1}.$$

Moreover, it is beyond a doubt that $\mathcal{G}'' \uparrow$ is a filterbase on $X \uparrow$ inside $\mathbb{V}^{(B)}$. By construction, $\mathcal{G}' \supset \mathcal{G}''$ and so $\mathcal{G}' \uparrow \downarrow \supset \mathcal{G}'' \uparrow \downarrow$. Therefore, $\llbracket \mathcal{G}' \uparrow \supset \mathcal{G}'' \uparrow \rrbracket = \mathbb{1}$, implying that $\llbracket \text{fil}\{\mathcal{G}' \uparrow\} \supset \text{fil}\{\mathcal{G}'' \uparrow\} \rrbracket = \mathbb{1}$, with $\text{fil}\{\mathcal{B}\}$ standing as usual for the collection of supersets of the members of \mathcal{B} . Moreover,

$$\begin{aligned} &\llbracket (\forall F_1 \in \mathcal{G}' \uparrow)(\exists F_2 \in \mathcal{G}'' \uparrow)(F_1 \supset F_2) \rrbracket \\ &= \bigwedge_{F_1 \in \mathcal{G}' \uparrow} \llbracket (\exists F_2 \in \mathcal{G}'' \uparrow)(F_1 \supset F_2) \rrbracket = \mathbb{1} \end{aligned}$$

since $G_1 \uparrow \in \mathcal{G}' \uparrow$ for all $G_1 \in \mathcal{G}$ such that $\llbracket F_1 \supset G_1 \uparrow \rrbracket = \mathbb{1}$.

This yields $\llbracket \text{fil}\{\mathcal{G}' \uparrow\} \subset \text{fil}\{\mathcal{G}'' \uparrow\} \rrbracket = \mathbb{1}$ by the transfer principle for $\mathbb{V}^{(B)}$. \triangleright

4.8.6. The filter \mathcal{G}^\uparrow inside $\mathbb{V}^{(B)}$, we have constructed in 4.8.5, is the *ascent* of \mathcal{G} .

4.8.7. Let \mathcal{G} be a filterbase on the descent $X\downarrow$ of a nonempty X inside $\mathbb{V}^{(B)}$. Let further $\text{mix}(\mathcal{G})$ stand for the collection of mixings of nonempty families of members of \mathcal{G} . If \mathcal{G} consists of cyclic sets then $\text{mix}(\mathcal{G})$ is a filterbase on $X\downarrow$ and $\text{mix}(\mathcal{G}) \supset \mathcal{G}$. Moreover, $\mathcal{G}^\uparrow = \text{mix}(\mathcal{G})^\uparrow$.

◁ Take $U, V \in \text{mix}(\mathcal{G})$. This means that there exist sets Ξ and H , partitions of unity $(b_\xi)_{\xi \in \Xi}$ and $(c_\eta)_{\eta \in H}$, and families $(U_\xi)_{\xi \in \Xi}$ and $(V_\eta)_{\eta \in H}$ of members of \mathcal{G} such that $b_\xi U = b_\xi U_\xi$ for all $\xi \in \Xi$ and $c_\eta V = c_\eta V_\eta$ for all $\eta \in H$.

Let $W_{(\xi, \eta)} \subset U_\xi \cap V_\eta$ be some element of \mathcal{G} . Put $d_{(\xi, \eta)} := b_\xi \wedge c_\eta$. Clearly, $(d_{(\xi, \eta)})_{(\xi, \eta) \in \Xi \times H}$ is a partition of unity.

Consider the set $W := \sum_{(\xi, \eta) \in \Xi \times H} d_{(\xi, \eta)} W_{(\xi, \eta)}$, i.e. the collection of the corresponding mixings of members of $W_{(\xi, \eta)}$. Obviously, $d_{(\xi, \eta)} U = b_\xi c_\eta U = c_\eta b_\xi U_\xi \supset d_{(\xi, \eta)} W_{(\xi, \eta)}$ and, by the same reasons, $d_{(\xi, \eta)} V \supset d_{(\xi, \eta)} W_{(\xi, \eta)}$. Consequently, $W \subset U \cap V$ and $W \in \text{mix}(\mathcal{G})$.

Since \mathcal{G} consists of cyclic sets; therefore, considering 4.8.2 and 4.8.3, we infer that $\text{mix}(\mathcal{G})' = \text{mix}(\mathcal{G}')$, so completing the proof. ▷

4.8.8. Given a filter \mathcal{F} on X inside $\mathbb{V}^{(B)}$, put $\mathcal{F}^\downarrow := \text{fil}\{F\downarrow : F \in \mathcal{F}\}$. The filter \mathcal{F}^\downarrow on $X\downarrow$ is called the *descent* of \mathcal{F} . A filterbase \mathcal{G} on $X\downarrow$ is *extensional* provided that there is a filter \mathcal{F} on X such that $\text{fil}\{\mathcal{G}\} = \mathcal{F}^\downarrow$. A filterbase \mathcal{G} on $X\downarrow$ is *cyclic* provided that $\text{fil}\{\mathcal{G}\}$ has a base consisting of cyclic sets. (Note that the epithet “cyclic” sometimes replaces “extensional” in this context in the literature.)

4.8.9. A filter \mathcal{F} is extensional if and only if \mathcal{F} is a cyclic filter and $\mathcal{F} = \text{fil}\{\text{mix}(\mathcal{F})\}$.

◁ Everything follows from 4.8.2, 4.8.3, and 4.8.7. ▷

4.8.10. If \mathcal{F} and \mathcal{G} are extensional filters on the same set then $\mathcal{F} \supset \mathcal{G} \leftrightarrow [\mathcal{F}^\uparrow \supset \mathcal{G}^\uparrow] = \mathbf{1}$.

◁ If $\mathcal{F} \supset \mathcal{G}$ then $\mathcal{F}' \supset \mathcal{G}'$ and so $[\mathcal{F}^\uparrow \supset \mathcal{G}^\uparrow] = \mathbf{1}$. Therefore, $\mathcal{F}^\uparrow \downarrow \supset \mathcal{G}^\uparrow \downarrow$, i.e. $\mathcal{F}^\uparrow \downarrow \supset \mathcal{G}^\uparrow \downarrow$. It suffices to recall 4.8.8. ▷

4.8.11. A *proultrafilter* is a maximal element of the set of extensional filters.

4.8.12. Each *proultrafilter* is a maximal element of the set of cyclic filters.

◁ If \mathcal{A} is proultrafilter and \mathcal{F} is a cyclic filter finer than \mathcal{A} then $\mathcal{A} \subset \mathcal{F} \subset \text{mix}(\mathcal{F})$. Hence, $\mathcal{A} = \mathcal{F}$. Conversely, assume that \mathcal{A} is a maximal cyclic filter. Then $\mathcal{A} = \text{mix}(\mathcal{A})$ and so \mathcal{A} is a proultrafilter. ▷

4.8.13. Each *proultrafilter* on $X\downarrow$ is the descent of an ultrafilter on X .

◁ This is immediate from 4.8.8. ▷

4.8.14. *The following hold:*

(1) *If $f : X \rightarrow Y$ inside $\mathbb{V}^{(B)}$ and $[\mathcal{F}$ is a filter on X] = $\mathbb{1}$ then*

$$f(\mathcal{F})^\downarrow = f\downarrow(\mathcal{F}^\downarrow);$$

(2) *If $f : X\downarrow \rightarrow Y\downarrow$ is an extensional mapping and \mathcal{F} is a filter on $X\downarrow$ then*

$$f(\mathcal{F})^\uparrow = f\uparrow(\mathcal{F}^\uparrow);$$

(3) *The image of an extensional filter under an extensional mapping is extensional;*

(4) *The image of a poultrafilter under an extensional mapping is a poultrafilter.*

◁ (1): Using the definitions and properties of the descent $f\downarrow$ of f , observe that

$$\begin{aligned} G \in f(\mathcal{F})^\downarrow &\leftrightarrow (\exists U \in f(\mathcal{F})\downarrow)(G \supset U\downarrow) \leftrightarrow (\exists F \in \mathcal{F}\downarrow)(G \supset f(F)\downarrow) \\ &\leftrightarrow (\exists F \in \mathcal{F}\downarrow)(G \supset f\downarrow(F\downarrow)) \leftrightarrow (\exists F \in \mathcal{F}^\downarrow)(G \supset f\downarrow(F)) \leftrightarrow G \in f\downarrow(\mathcal{F}^\downarrow). \end{aligned}$$

(2): Using the properties of the ascent $f\uparrow$ of f , estimate the truth values as follows:

$$\begin{aligned} \llbracket G \in f\uparrow(\mathcal{F}^\uparrow) \rrbracket &= \llbracket (\exists U \in f\uparrow(\mathcal{F}^\uparrow))(G \supset U) \rrbracket \\ &= \llbracket (\exists F \in \mathcal{F}^\uparrow)(G \supset f\uparrow(F)) \rrbracket = \bigvee_{F \in \mathcal{F}} \llbracket G \supset f\uparrow(F) \rrbracket \\ &= \bigvee_{F \in \mathcal{F}} \llbracket G \supset f(F)\uparrow \rrbracket = \bigvee_{U \in f(\mathcal{F})'} \llbracket G \supset U \rrbracket = \llbracket (\exists U \in f(\mathcal{F})'\uparrow)(G \supset U) \rrbracket \\ &= \llbracket (\exists U \in f(\mathcal{F})^\uparrow)(G \supset U) \rrbracket = \llbracket G \in f(\mathcal{F})^\uparrow \rrbracket. \end{aligned}$$

(3): Using (2) and (1) successively, obtain

$$f(\mathcal{F})^{\uparrow\downarrow} = f\uparrow(\mathcal{F}^\uparrow)^\downarrow = f\uparrow\downarrow(\mathcal{F}^{\uparrow\downarrow}) = f(\mathcal{F}^{\uparrow\downarrow}).$$

The last equality ensures the claim.

(4): If $f : X\downarrow \rightarrow Y\downarrow$ is an extensional mapping and \mathcal{F} is a poultrafilter then \mathcal{F}^\uparrow is an ultrafilter on X inside $\mathbb{V}^{(B)}$. Consequently, $f\uparrow(\mathcal{F}^\uparrow)$ is an ultrafilter on Y inside $\mathbb{V}^{(B)}$. Therefore, $f\uparrow(\mathcal{F}^\uparrow)^\downarrow$ is a poultrafilter. It suffices to note that $f\uparrow(\mathcal{F}^\uparrow)^\downarrow = f(\mathcal{F}^{\uparrow\downarrow}) = f(\mathcal{F})$ by (3). ▷

4.9. Essential and Proideal Points of Cyclic Monads

In this section we give a test for a filter to be cyclic and introduce some relevant concepts we need in the sequel.

We start with a standard complete Boolean algebra B and the corresponding Boolean-valued universe $\mathbf{V}^{(B)}$ which is thought of as composed of internal sets.

4.9.1. If A is an external set then the *cyclic hull* $\text{mix}(A)$ of A is introduced as follows:

An element $x \in \mathbf{V}^{(B)}$ belongs to $\text{mix}(A)$ if there are an internal family $(a_\xi)_{\xi \in \Xi}$ of elements of A and an internal partition $(b_\xi)_{\xi \in \Xi}$ of unity in B such that x is the mixing of $(a_\xi)_{\xi \in \Xi}$ by $(b_\xi)_{\xi \in \Xi}$, i.e. $b_\xi x = b_\xi a_\xi$ for $\xi \in \Xi$ or, equivalently, $x = \text{mix}_{\xi \in \Xi}(b_\xi a_\xi)$.

The monad $\mu(\mathcal{F})$ of a filter \mathcal{F} is *cyclic* if $\mu(\mathcal{F})$ coincides with its cyclic hull $\text{mix}(\mu(\mathcal{F}))$.

4.9.2. Theorem. *A standard filter is cyclic if and only if so is its monad.*

$\triangleleft \rightarrow$: Let \mathcal{F} be a standard cyclic filter. Take an internal set Ξ , an internal partition of unity $(b_\xi)_{\xi \in \Xi}$, and a family $(x_\xi)_{\xi \in \Xi}$ of points of $\mu(\mathcal{F})$. By hypothesis, \mathcal{F} has some base \mathcal{G} consisting of cyclic sets. Hence, $\mu(\mathcal{F}) = \bigcap \{G : G \in {}^\circ\mathcal{G}\}$. If x is a mixing of $(x_\xi)_{\xi \in \Xi}$ by $(b_\xi)_{\xi \in \Xi}$ then x belongs to every standard G in \mathcal{G} (since $x_\xi \in G$ for $\xi \in \Xi$). Consequently, $\mu(\mathcal{F}) \supset \text{mix}(\mu(\mathcal{F})) \supset \mu(\mathcal{F})$.

\leftarrow : The monad $\mu(\mathcal{F})$ is now a cyclic external set. Taking an infinitesimal member F in \mathcal{F} (i.e. $F \subset \mu(\mathcal{F})$), note that $F_0 := \text{mix}(F) \subset \text{mix}(\mu(\mathcal{F})) \subset \mu(\mathcal{F})$. Therefore, the internal set F_0 is infinitesimal and lies in \mathcal{F} . So, $(\forall^{\text{st}} F \in \mathcal{F})(\exists F_0 \in \mathcal{F})(F_0 = \text{mix}(F_0) \wedge F \supset F_0)$. By transfer, \mathcal{F} has a cyclic base. \triangleright

4.9.3. Theorem. *Let \mathcal{F} be a standard filter on $X \downarrow$. Put*

$$\mathcal{F} \uparrow \downarrow := \text{fil}\{F \uparrow \downarrow : F \in \mathcal{F}\}.$$

Then $\text{mix}(\mu(\mathcal{F})) = \mu(\mathcal{F} \uparrow \downarrow)$, and $\mathcal{F} \uparrow \downarrow$ is the finest cyclic filter coarser than \mathcal{F} .

\triangleleft Clearly, $\mathcal{F} \uparrow \downarrow \subset \mathcal{F}$. By Theorem 4.9.2 $\mu(\mathcal{F} \uparrow \downarrow) \supset \mu(\mathcal{F})$ and $\mu(\mathcal{F} \uparrow \downarrow) \supset \text{mix}(\mu(\mathcal{F}))$. Take $x \in \mu(\mathcal{F} \uparrow \downarrow)$. Using the definition of monad and the properties of mixing, obtain

$$(\forall^{\text{st}} F \in \mathcal{F})(\exists (b_\xi)_{\xi \in \Xi})(\exists (x_\xi)_{\xi \in \Xi})(\forall \xi \in \Xi)(x_\xi \in F \wedge b_\xi x = b_\xi x_\xi).$$

Therefore,

$$\begin{aligned} & (\forall^{\text{st fin}} \mathcal{F}_0 \subset \mathcal{F})(\exists (b_\xi)_{\xi \in \Xi})(\exists (x_\xi)_{\xi \in \Xi})(\forall F_0 \in \mathcal{F}_0) \\ & (\forall \xi \in \Xi)(x_\xi \in F \wedge b_\xi x_\xi = b_\xi x). \end{aligned}$$

By idealization

$$(\exists(b_\xi)_{\xi \in \Xi})(\exists(x_\xi)_{\xi \in \Xi})(\forall^{\text{st}} F \in \mathcal{F})(\forall \xi \in \Xi)(x_\xi F \wedge b_\xi x_\xi = b_\xi x).$$

This means that there are some elements $(x_\xi)_{\xi \in \Xi}$ in $\mu(\mathcal{F})$ such that $x = \sum_{\xi \in \Xi} b_\xi x_\xi$, i.e. $x \in \text{mix}(\mu(\mathcal{F}))$. Conclude that $\mu(\mathcal{F}\uparrow\downarrow) = \text{mix}(\mu(\mathcal{F}))$.

Now, let \mathcal{G} be a cyclic filter satisfying $\mathcal{G} \subset \mathcal{F}$. Therefore, $\text{mix}(\mu(\mathcal{G})) = \mu(\mathcal{G}) \supset \text{mix}(\mu(\mathcal{F})) = \mu(\mathcal{F}\uparrow\downarrow)$. Consequently, $\mathcal{G} \subset \mathcal{F}\uparrow\downarrow$. \triangleright

4.9.4. Let x be an internal point of $X\downarrow$. Define the standard filter (x) on $X\downarrow$ as

$$(x) := * \{U \subset X\downarrow : x \in U\}.$$

In other words, (x) comprises exactly those standard subsets of $X\downarrow$ that contain x . A member x of $X\downarrow$ is an *essential point* of $X\downarrow$ (in symbols, $x \in e(X)$) provided that $(x)\uparrow\downarrow$ is a proultrafilter on $X\downarrow$.

4.9.5. Every point x of the monad of a standard proultrafilter \mathcal{F} is essential. Moreover,

$$\mathcal{F} = (x)\uparrow\downarrow = (x)\uparrow\downarrow = \text{fil}\{*\{U\uparrow\downarrow : x \in U \wedge U \subset X\downarrow\}\}.$$

\triangleleft Since $\mu(\mathcal{F})$ and (x) is an ultramonad, i.e., the monad of an ultrafilter; therefore, $(x) \supset \mathcal{F}$. Consequently, $(x)\uparrow\downarrow \supset \mathcal{F}\uparrow\downarrow = \mathcal{F}$. By 4.8.12 $\mathcal{F} = (x)\uparrow\downarrow$. From 4.8.5 it follows that $(x)\uparrow\downarrow^\uparrow = (x)^\uparrow$. By 4.8.13 x is an essential point. Finally, $(x)\uparrow\downarrow = \mathcal{F}\uparrow\downarrow = \mathcal{F} = (x)\uparrow\downarrow$. \triangleright

4.9.6. The image of an essential point under an extensional mapping is an essential point of the target space.

\triangleleft Assume that x is an essential point of $X\downarrow$ and $f : X\downarrow \rightarrow Y\downarrow$ is an extensional mapping. There is a proultrafilter \mathcal{F} satisfying $x \in \mu(\mathcal{F})$. Clearly, $f(x) \in f(\mu(\mathcal{F})) = \mu(f(\mathcal{F}))$. Indeed, by idealization we have

$$\begin{aligned} y \in \mu(f(\mathcal{F})) &\leftrightarrow (\forall^{\text{st}} F \in \mathcal{F})(y \in f(F)) \\ &\leftrightarrow (\forall^{\text{st fin}} \mathcal{F}_0 \subset \mathcal{F})(\exists x)(\forall F \in \mathcal{F}_0)(x \in F \wedge y = f(x)) \\ &\leftrightarrow (\exists x)(\forall^{\text{st}} F \in \mathcal{F} x \in F \wedge y = f(x)) \leftrightarrow (\exists x \in \mu(\mathcal{F}))(y = f(x)) \\ &\leftrightarrow y \in f(\mu(\mathcal{F})). \end{aligned}$$

We are done on recalling 4.8.14. \triangleright

4.9.7. Assume that E is a standard set and X is a standard member of $\mathbb{V}^{(B)}$. Consider the product X^{E^\wedge} inside $\mathbb{V}^{(B)}$, with E^\wedge the standard name of E in $\mathbb{V}^{(B)}$. If x is an essential point of $X^{E^\wedge}\downarrow$ then $x\downarrow(e)$ is an essential point of $X\downarrow$ for all standard $e \in E$.

◁ Since $x \in X^{E^\wedge} \downarrow$; therefore, $\llbracket x : E^\wedge \rightarrow X \rrbracket = \mathbb{1}$, i.e. $x \downarrow : E \rightarrow X \downarrow$ and $\llbracket x \downarrow(e) = x(e^\wedge) \rrbracket = \mathbb{1}$ for all $e \in E$ by the definition of the descent $x \downarrow$.

Given a standard $e \in E$, consider the mapping sending $x \in X^{E^\wedge} \downarrow$ to the point $x(e^\wedge)$ in $X \downarrow$. For $x_1, x_2 \in X^{E^\wedge} \downarrow$ it is clear that

$$\begin{aligned} \llbracket x_1 = x_2 \rrbracket &= \llbracket (\forall e \in E^\wedge)(x_1(e) = x_2(e)) \rrbracket = \bigwedge_{e \in E} \llbracket x_1(e^\wedge) = x_2(e^\wedge) \rrbracket \\ &\leq \llbracket x_1(e^\wedge) = x_2(e^\wedge) \rrbracket, \end{aligned}$$

i.e. the above mapping is extensional. By 4.9.6 $x(e^\wedge)$ is an essential point of $X \downarrow$. We are done on recalling that $x \downarrow(e) = x(e^\wedge)$ by the definition of descent. ▷

4.9.8. Let \mathcal{F} be a cyclic filter on $X \downarrow$, and let ${}^e\mu(\mathcal{F}) := \mu(\mathcal{F}) \cap e(X)$ stand for the set of essential points of the monad of \mathcal{F} . Then ${}^e\mu(\mathcal{F}) = {}^e\mu(\mathcal{F}^{\uparrow\downarrow})$.

◁ Take $x \in {}^e\mu(\mathcal{F})$. By definition, x belongs to the monad of some proultrafilter \mathcal{G} . Hence, $\mu(\mathcal{G}) \cap \mu(\mathcal{F}) \neq \emptyset$ and so $\mathcal{G} \supset \mathcal{F}$. By 4.8.10 $\mathcal{G}^{\uparrow\downarrow} \supset \mathcal{F}^{\uparrow\downarrow}$ and $x \in \mu(\mathcal{G}) \subset \mu(\mathcal{F}^{\uparrow\downarrow})$.

If we now assume that $x \in {}^e\mu(\mathcal{F}^{\uparrow\downarrow})$ then there is an ultrafilter \mathcal{G} on X inside $\mathbb{V}^{(B)}$ such that $x \in \mu(\mathcal{G}^\downarrow)$ and $\mathcal{G} \supset \mathcal{F}^\uparrow$. Since $\mathcal{F} = \mathcal{F}^{\uparrow\downarrow} \subset \mathcal{F}^{\uparrow\downarrow} \subset \mathcal{G}^\downarrow$ by 4.9.7; therefore, $\mu(\mathcal{F}) \supset \mu(\mathcal{G}^\downarrow)$. Consequently, $x \in {}^e\mu(\mathcal{F})$. ▷

4.9.9. Let A be a subset of the descent $X \downarrow$ in question. The set $(X - A^\uparrow) \downarrow$, denoted by A^c , is the *procomplement* or *cyclic complement* of A .

A point x in $X \downarrow$ is *proideal* provided that x belongs to the procomplement of each standard finite subset of $X \downarrow$. The symbol $p(X)$ stands for the set of proideal points of $X \downarrow$.

4.9.10. If $X \downarrow$ lacks proideal points then X is a finite set inside $\mathbb{V}^{(B)}$.

◁ By idealization, there is a standard finite set Y in $X \downarrow$ satisfying $Y^e = \emptyset$. Hence, $\llbracket X - Y^\uparrow = \emptyset^\uparrow \rrbracket = \mathbb{1}$, i.e. $X = Y^\uparrow$. ▷

4.9.11. If X is an infinite set inside $\mathbb{V}^{(B)}$ then the proideal points of $X \downarrow$ comprise a cyclic monad. The descent of the cyclic filter with monad $p(X)$ is the cofinite filter on X inside $\mathbb{V}^{(B)}$.

◁ The procomplements of finite subsets of $X \downarrow$ comprise a filterbase. Indeed, from $(Y \cup Z)^{\uparrow\downarrow} \supset Y^{\uparrow\downarrow} \cup Z^{\uparrow\downarrow}$ it follows that $(Y \cup Z)^\uparrow \supset Y^\uparrow \cup Z^\uparrow$ and $\llbracket X - (Y \cup Z)^\uparrow \subset X - (Y^\uparrow \cup Z^\uparrow) \rrbracket = \mathbb{1}$. Hence, $(Y \cup Z)^c \subset (X - Y^\uparrow) \downarrow \cap (X - Z^\uparrow) \downarrow = Y^c \cap Z^c$. By Theorem 4.9.2 $p(X)$ is a cyclic monad. Let ${}_p\mathcal{F}$ stand for the filter with monad $p(X)$, i.e. the filter of procomplements of finite subsets of $X \downarrow$. Assume further that ${}_{cf}\mathcal{F}(X)$ is the filter of cofinite subsets of X inside $\mathbb{V}^{(B)}$ which is the

cofinite filter on X by definition. By 4.9.4

$$\begin{aligned} \llbracket Y \in_{\text{cf}} \mathcal{F}(X) \rrbracket &= \llbracket (\exists Z \in \mathcal{P}_{\text{fin}}(X))(Y \supset X - Z) \rrbracket \\ &= \bigvee_{A \in \mathcal{P}_{\text{fin}}(X \downarrow)} \llbracket Y \supset X - A \uparrow \rrbracket = \bigvee_{A \in \mathcal{P}_{\text{fin}}(X \downarrow)} \llbracket Y \supset A^c \uparrow \rrbracket \\ &= \bigvee_{Z \in_p \mathcal{F}} \llbracket Y \supset Z \uparrow \rrbracket = \llbracket Y \in_p \mathcal{F} \uparrow \rrbracket. \end{aligned}$$

Consequently, $_{\text{cf}} \mathcal{F}(X) = {}_p \mathcal{F} \uparrow$. \triangleright

4.10. Descending Compact and Precompact Spaces

In this section we apply cyclic monads to describing the descents of topological spaces inside Boolean valued models of set theory. The results below follow the ideas of the classical articles by Robinson [421] and Luxemburg [328].

For the sake of simplicity, we always consider a nonempty uniform space (X, \mathcal{U}) inside $\mathbb{V}^{(B)}$. We also assume the standard environment: using nonstandard tools, we consider B, X, \mathcal{U} , etc. standard sets. As usual, we write $x \approx y$ whenever $(x, y) \in \mu(\mathcal{U} \downarrow)$.

4.10.1. The uniform space $(X \downarrow, \mathcal{U} \downarrow)$ is *procompact* or *cyclically compact* provided that (X, \mathcal{U}) is compact inside $\mathbb{V}^{(B)}$. A similar sense resides in the notion of *pro-total-boundedness* and so on. The terms like “cyclic compactness” may be encountered in several publications.

4.10.2. Theorem. *For a standard space X the following are equivalent:*

- (1) $X \downarrow$ is procompact;
- (2) Every essential point of $X \downarrow$ is nearstandard;
- (3) Every essential ideal point of $X \downarrow$ is nearstandard.

\triangleleft (1) \rightarrow (2): Let x be an essential point of $X \downarrow$. Then x belongs to the monad of the poultrafilter $(x) \uparrow \downarrow$. It is thus true inside $\mathbb{V}^{(B)}$ that $(x) \uparrow$ converges to some member y of X . By the maximum and transfer principles, there is a standard point y in $X \downarrow$ satisfying $(x) \uparrow \downarrow \supset \mathcal{U} \uparrow(y)$. Hence, $\mu((x) \uparrow \downarrow) \subset \mathcal{U} \downarrow(y)$ and so $x \approx y$. In other words, x is a nearstandard point.

(2) \rightarrow (3): Obvious.

(3) \rightarrow (1): We have to demonstrate inside $\mathbb{V}^{(B)}$ that each ultrafilter on X is an adherent point.

Without loss of generality, assume that \mathcal{F} is not a principal ultrafilter. In other words, \mathcal{F} is finer than the cofinite filter on X inside $\mathbb{V}^{(B)}$. From 4.9.6 it follows that $\mu(\mathcal{F} \downarrow) \subset p(X)$. If $x \in \mu(\mathcal{F} \downarrow)$ then $\mathcal{F} = (x) \uparrow$ by 4.9.8. Moreover, x is an essential point. By hypothesis such a point is nearstandard, i.e. there is a standard point y in $X \downarrow$ satisfying $\mathcal{U} \downarrow(y) \cap \mu(\mathcal{F} \downarrow) \neq \emptyset$. This implies that y is an adherent point of \mathcal{F} inside $\mathbb{V}^{(B)}$. \triangleright

4.10.3. It is easy from 4.10.2 that the Boolean valued test for procompactness differs from its classical analog: “A compact space is a space consisting of nearstandard points.” Procompact but not compact spaces are galore, providing a plenty of inessential points.

We note also that the joint application of 4.10.2 and 4.9.7 allows us easily to produce a nonstandard proof for a natural analog of Tychonoff’s Theorem for a product of procompact spaces, the “descent of Tychonoff’s Theorem in $\mathbb{V}^{(B)}$.”

4.10.4. *A standard space is the descent of a totally bounded uniform space if and only if its every essential point is prenearstandard.*

$\triangleleft \rightarrow$: Let x be an essential point of $X \downarrow$. Then $(x)^\uparrow$ is an ultrafilter inside $\mathbb{V}^{(B)}$ and so $(x)^\uparrow$ is a Cauchy filter on X for X is totally bounded inside $\mathbb{V}^{(B)}$. The descent of a Cauchy filter on X is a Cauchy filter on the descent of X . Therefore, x belongs to the monad of a Cauchy filter and so x is prenearstandard.

\leftarrow : Take an ultrafilter \mathcal{F} on X inside $\mathbb{V}^{(B)}$. The task is to show that \mathcal{F} is the Cauchy filter inside $\mathbb{V}^{(B)}$. Take a point x from the monad of the descent \mathcal{F}^\downarrow of \mathcal{F} . Then x is essential and so x is prenearstandard. Consequently, the microhalo $\mathcal{U}^\downarrow(x)$ of x is the monad of a Cauchy filter. It follows that \mathcal{F}^\downarrow is a Cauchy filter. \triangleright

4.11. Proultrafilters and Extensional Filters

In 4.8 we have applied the monadology of infinitesimal analysis to study the cyclic filters characteristic of Boolean valued topology. In this section we give some tests for the monads of proultrafilters and extensional filters and discuss a few relevant properties. Throughout this section we fix a complete Boolean algebra B and the corresponding separated Boolean valued universe $\mathbb{V}^{(B)}$.

4.11.1. Let X be a cyclic set which is the descent of some member of $\mathbb{V}^{(B)}$. By μ_d we denote the taking of the (discrete) *monadic hull*. In other words, $\mu_d(\emptyset) := \emptyset$ and, in case U is a nonempty subset of X , we assume that $\mu_d(U)$ is the standardization of the external filter of supersets of U :

$$x \in \mu_d(U) \leftrightarrow ((\forall^{\text{st}} V \subset X) U \subset V \rightarrow x \in U).$$

By analogy we define the *cyclic monadic hull* μ_c as follows:

$$x \in \mu_c(U) \leftrightarrow (\forall^{\text{st}} V)(V = V \uparrow \downarrow \wedge V \subset X \wedge U \subset V \rightarrow x \in V).$$

In other words, if U is nonempty then the cyclic monadic hull $\mu_c(U)$ of U is the monad of the cyclic hull of the standardization of the filter of supersets of U .

4.11.2. The cyclic monadic hull of a set U is the cyclic hull of the monadic hull of U ; in symbols,

$$\mu_c(U) = \text{mix}(\mu_d(U)).$$

◁ Assume that $U \neq \emptyset$ and U is a standard set satisfying $V \supset \text{mix}(\mu_d(U))$. Using 4.9.3, note that $V \supset W \uparrow \downarrow$ for some member W of $^*\{Y \subset X : Y \supset U\}$. Consequently, $V \supset \mu_c(U)$. It follows that $\mu_c(U) \subset \text{mix}(\mu_d(U))$ since the rightmost set is a monad.

Conversely, if $V \supset \mu_c(U)$ and V is standard then V includes the cyclic hull of a superset of U and so $V \supset U$. Hence $V \supset \mu(^*\{W : W \supset U\}) \uparrow \downarrow$ and we are done on recalling 4.9.3. ▷

4.11.3. Since X is cyclic, X is the descent of the ascent of X . Consequently, each proultrafilter on X is the descent of some ultrafilter on the ascent $X \uparrow$ of X .

4.11.4. Theorem. A filter \mathcal{F} is a proultrafilter on X if and only if the monad $\mu(\mathcal{F})$ of \mathcal{F} is cyclic and $\mu(\mathcal{F})$ is easy to catch by each cyclic subset of X : either $\mu(\mathcal{F}) \subset U$ or $\mu(\mathcal{F}) \subset X - U$ for every cyclic set U in X .

◁ →: We are to prove the following:

(\mathcal{F} is a proultrafilter)

$$\leftrightarrow \mu(\mathcal{F}) = \text{mix}(\mu(\mathcal{F})) \wedge (\forall^{\text{st}} V)(V = V \uparrow \downarrow \rightarrow \mu(\mathcal{F}) \subset V \vee \mu(\mathcal{F}) \subset V').$$

Given a standard subset V of X , note that $\mu(\mathcal{F}) \cap V = \emptyset$, or $\mu(\mathcal{F}) \cap V \neq \emptyset$. In the first case $V' := X - V \in \mathcal{F}$. In the second case we observe the filter \mathcal{G} whose monad is $\mu(\mathcal{F}) \cap V$. Clearly, if \mathcal{F} is a proultrafilter and V is a cyclic set then $\mathcal{G} = \mathcal{F}$ by Theorem 4.9.2. Hence, $V \in \mathcal{F}$.

←: Take some cyclic filter \mathcal{G} finer than \mathcal{F} . Obviously, $G \in \mathcal{G} \rightarrow G' \notin \mathcal{F}$ (otherwise we would have $G' \supset \mu(\mathcal{F}) \supset \mu(\mathcal{G})$). Hence, $G \in \mathcal{F}$, implying that $\mathcal{G} = \mathcal{F}$. ▷

4.11.5. Theorem. Let \mathcal{F} be a cyclic filter on X . The following are equivalent:

- (1) \mathcal{F} is a proultrafilter;
- (2) If \mathcal{E} is a finite set of subsets of X then either $(\bigcup \mathcal{E})' \in \mathcal{F}$ or $E \uparrow \downarrow \in \mathcal{F}$ for some $E \in \mathcal{E}$;
- (3) If \mathcal{E} is a finite family of cyclic subsets of X then \mathcal{F} contains either a member of \mathcal{E} or the complement of each member of \mathcal{E} ;
- (4) If U is an arbitrary set then either $U \uparrow \downarrow \in \mathcal{F}$ or $U' \in \mathcal{F}$;
- (5) If V is an arbitrary cyclic set then either $V \in \mathcal{F}$ or $V' \in \mathcal{F}$.

◁ To prove the implication (1) → (2), use the transfer principle and the test of Theorem 4.11.4.

Assume that \mathcal{F} is a standard filter and \mathcal{E} is a standard finite set of standard subsets of X . The two cases are possible: $\mu(\mathcal{F}) \cap \bigcup \mathcal{E} = \emptyset$ or $\mu(\mathcal{F}) \cap \bigcup \mathcal{E} \neq \emptyset$. In the first case it is obvious that $(\bigcup \mathcal{E})'$ belongs to \mathcal{F} . In the second case there is some E in \mathcal{E} satisfying $E \cap \mu(\mathcal{F}) \neq \emptyset$. Therefore, $E\uparrow\downarrow \cap \mu(\mathcal{F}) \neq \emptyset$. Since $E\uparrow\downarrow$ is standard, by Theorem 4.11.4 conclude that $E\uparrow\downarrow \supset \mu(\mathcal{F})$ and so $E\uparrow\downarrow \in \mathcal{F}$.

The implications (2) \rightarrow (3) \rightarrow (4) \rightarrow (5) raise no doubts. The implication (5) \rightarrow (1) ensues from Theorem 4.11.4 by transfer. \triangleright

4.11.6. Let \mathcal{F} be a filter on X . The filter $\mathcal{F}\uparrow\downarrow$ is a proultrafilter if and only if for each subset U of X either $U\uparrow\downarrow \in \mathcal{F}$ or there is some F in \mathcal{F} satisfying $F\uparrow\downarrow \subset U'$.

4.11.7. Let \mathcal{F} be a filter on X . For \mathcal{F} to be a proultrafilter on X it is necessary and sufficient that $\mathcal{F} = (\ddot{\mathcal{F}})\uparrow\downarrow$, where $\ddot{\mathcal{F}}$ is the grill of \mathcal{F} , i.e.

$$U \in \ddot{\mathcal{F}} \leftrightarrow (\forall F \in \mathcal{F})(U \cap F \neq \emptyset).$$

\triangleleft Assume that \mathcal{F} is a proultrafilter. Clearly, $\mathcal{F} \subset \ddot{\mathcal{F}}$ and so $\mathcal{F} = \mathcal{F}\uparrow\downarrow \subset (\ddot{\mathcal{F}})\uparrow\downarrow$. If $V \in (\ddot{\mathcal{F}})\uparrow\downarrow$ then $V \supset U\uparrow\downarrow$ for some U in $\ddot{\mathcal{F}}$. Moreover, $U\uparrow\downarrow$ belongs to \mathcal{F} by (4). Consequently, $V \in \mathcal{F}$.

Assume now that $\mathcal{F} = (\ddot{\mathcal{F}})\uparrow\downarrow$. Since each member of the right side is a superset of a cyclic set by definition; therefore, \mathcal{F} is a cyclic filter.

Let U be an arbitrary cyclic set. If $U \cap F = \emptyset$ for some $F \in \mathcal{F}$ then $U' \in \mathcal{F}$. If $U \cap F \neq \emptyset$ for all $F \in \mathcal{F}$ then U belongs to $(\ddot{\mathcal{F}})\uparrow\downarrow$ and so $U \in \mathcal{F}$. From (5) it follows that \mathcal{F} is a proultrafilter. \triangleright

4.11.8. The family $(\ddot{\mathcal{F}})\uparrow\downarrow$ in 4.11.7 is the *cyclic grill* or (rarely) *progrill* of \mathcal{F} .

The meaning of this concept as well as the way of its application becomes clear in regard to the ascending and descending technique of Boolean valued analysis. Recall that if \mathcal{E} is a family of nonempty subsets of $X\uparrow$ inside $\mathbb{V}^{(B)}$ then the *descent* $\mathcal{E}\downarrow$ of \mathcal{E} is defined by the formula

$$U \in \mathcal{E}\downarrow \leftrightarrow U \subset X \wedge (\exists E \in \mathcal{E}\downarrow)(U \supset E\downarrow).$$

4.11.9. Let \mathcal{F} be a filter with grill $\ddot{\mathcal{F}}$ inside $\mathbb{V}^{(B)}$. Then

$$(\ddot{\mathcal{F}})\downarrow = (\ddot{\mathcal{F}}\downarrow)\uparrow\downarrow.$$

\triangleleft Using the rules for calculating truth values inside $\mathbb{V}^{(B)}$, note that

$$\llbracket U\uparrow \in \ddot{\mathcal{F}} \rrbracket = \bigwedge_{F \in \mathcal{F}\downarrow} \llbracket A\uparrow \cap F\uparrow \neq \emptyset \rrbracket,$$

where A is a subset of X . \triangleright

4.11.10. *An extensional filter \mathcal{F} is a poultrafilter if and only if the cyclic grill of \mathcal{F} is a filter.*

◁ Clearly, \mathcal{F} is a poultrafilter if and only if the descent \mathcal{F}^\uparrow coincides with its grill inside $\mathbb{V}^{(B)}$. The last happens if and only if the grill of \mathcal{F}^\uparrow is a filter inside $\mathbb{V}^{(B)}$. We are done on appealing to 4.11.9. ▷

4.11.11. *A point x is essential if and only if x can be separated by a standard set from every standard cyclic set not containing x .*

◁ In symbols, the claim reads:

$$x \in {}^e X \leftrightarrow (\forall U = U\uparrow\downarrow)(x \notin U \rightarrow (\exists V = V\uparrow\downarrow) x \in V \wedge U \cap V = \emptyset).$$

Assume first that x is an essential point and U is a standard cyclic set satisfying $x \notin U$. By 4.11.3 the complement U' of U belongs to the filter $(x)\uparrow\downarrow$ generated by cyclic supersets of x (since $(x)\uparrow\downarrow$ is a poultrafilter by hypothesis). Hence, there is some V such that $x \in V$ and $V\uparrow\downarrow \cap U \neq \emptyset$.

Assume now that the separatedness condition is fulfilled. Then $(x)\uparrow\downarrow$ meets the conditions of Theorem 4.11.4. Indeed, let $U = U\uparrow\downarrow$ be an arbitrary cyclic set. We are to check that either U or U' belongs to $(x)\uparrow\downarrow$. If $x \in U$ then $U \in (x)\uparrow\downarrow$ by definition. If $x \in U'$ then $V \cap U = \emptyset$ for some $V \in (x)\uparrow\downarrow$ by hypothesis; i.e., $V \subset U'$ and $U' \in (x)\uparrow\downarrow$. ▷

4.11.12. *If the monad of an ultrafilter \mathcal{F} has an essential point then $\mu(\mathcal{F}) \subset {}^e X$ and, moreover, $\mathcal{F}\uparrow\downarrow$ is a poultrafilter.*

◁ Assume that V is an arbitrary cyclic set and $x \in \mu(\mathcal{F}) \cap {}^e X$. If $x \in V$ then $V \cap \mu(\mathcal{F}) \neq \emptyset$. Hence, $V \in \mathcal{F}$ and so $V \in \mathcal{F}\uparrow\downarrow$. If $x \notin V$ then $x \in U$ and $U \cap V = \emptyset$ for some cyclic set U by 4.11.11. Clearly, $U \in \mathcal{F}\uparrow\downarrow$. Therefore, $V' \in \mathcal{F}\uparrow\downarrow$. Recalling 4.11.5, conclude that $\mathcal{F}\uparrow\downarrow$ is a poultrafilter. As was mentioned, this implies $\mu(\mathcal{F})\uparrow\downarrow \subset {}^e X$. Since $\mu(\mathcal{F}\uparrow\downarrow) = \text{mix}(\mu(\mathcal{F}))$, we are done by 4.9.3. ▷

4.11.13. Theorem. *A filter \mathcal{F} is extensional if and only if the monad of \mathcal{F} is the cyclic monadic hull of its essential points.*

◁ In symbols, the claim reads:

$$(\mathcal{F} \text{ is extensional}) \leftrightarrow \mu(\mathcal{F}) = \text{mix}(\mu_d({}^e \mu(\mathcal{F}))).$$

The fact that \mathcal{F} is extensional may be rewritten as follows: $\llbracket \mathcal{F}^\uparrow \text{ is a filter on } X^\uparrow \rrbracket = \mathbb{1}$. By the Boolean valued transfer principle, there is some set \mathfrak{A} of poultrafilters on X such that

$$\llbracket F \in \mathcal{F}^\uparrow \rrbracket = \bigwedge_{\mathcal{A} \in \mathfrak{A}} \llbracket F \in \mathcal{A}^\uparrow \rrbracket.$$

Given a cyclic set F in X , we thus have

$$F \in \mathcal{F}^{\uparrow\downarrow} \leftrightarrow F \in \bigcap_{\mathcal{A} \in \mathfrak{A}} \mathcal{A}^{\uparrow\downarrow}.$$

If now F is cyclic then

$$F \supset \mu(\mathcal{F}^{\uparrow\downarrow}) \leftrightarrow F \supset \mu_d \left(\bigcup_{\mathcal{A} \in {}^\circ\mathfrak{A}} \mu(\mathcal{A}^{\uparrow\downarrow}) \right),$$

with ${}^\circ\mathfrak{A}$ the standard part of \mathfrak{A} , i.e. the external set of standard members of \mathfrak{A} . Using 4.11.2, note that

$$\mu(\mathcal{F}^{\uparrow\downarrow}) = \text{mix} \left(\mu \left(\bigcup_{\mathcal{A} \in {}^\circ\mathfrak{A}} \mu(\mathcal{A}^{\uparrow\downarrow}) \right) \right).$$

It suffices to recall that the monad of every poultrafilter consists of essential points by 4.9.5 and we may take as \mathfrak{A} the set of ultrafilters finer than \mathcal{F} . \triangleright

4.11.14. *A standard set U is cyclic if and only if U coincides with the cyclic monadic hull of the set of essential points of U .*

4.11.15. *Let \mathcal{F} be a filter on X , and let b be a member of the Boolean algebra B . Assume further that $b\mathcal{F}$ stands for the image of \mathcal{F} under multiplication by b . Then*

$$b(b\mathcal{F})^\uparrow = b\mathcal{F}^\uparrow.$$

\triangleleft Proceed by calculating the truth values

$$\begin{aligned} \llbracket (b\mathcal{F})^\uparrow = \mathcal{F}^\uparrow \rrbracket &\geq \bigwedge_{F \in \mathcal{F}} \llbracket (bF)^\uparrow \in \mathcal{F}^\uparrow \rrbracket \geq \bigwedge_{F \in \mathcal{F}} \llbracket (bF)^\uparrow = F^\uparrow \rrbracket \\ &\geq \bigwedge_{F \in \mathcal{F}} \bigwedge_{x \in X} \llbracket bx \in F^\uparrow \rrbracket \geq \bigwedge_{F \in \mathcal{F}} \bigwedge_{x \in X} \llbracket bx = x \rrbracket \geq b, \end{aligned}$$

which completes the proof. \triangleright

4.11.16. *Assume that \mathcal{F} and \mathcal{G} are filters on the same set inside $\mathbb{V}^{(B)}$ and $b \in B$. Then*

$$b\mathcal{F} = b\mathcal{G} \leftrightarrow b\mathcal{F}^\downarrow = b\mathcal{G}^\downarrow.$$

\triangleleft If $\llbracket \mathcal{F} \subset \mathcal{G} \rrbracket \geq b$ then to each $F \in \mathcal{F}^\downarrow$ there is some G in \mathcal{G}^\downarrow satisfying

$$\llbracket F^\uparrow \supset G^\uparrow \rrbracket = \llbracket \mathcal{F} \supset \mathcal{G} \rrbracket \geq b$$

by the maximum principle. In other words, $bF^\uparrow \supset bG^\uparrow$. Hence, $bF \supset bG$ for all cyclic F and G . Consequently, $b\mathcal{F}^\downarrow \subset b\mathcal{G}^\downarrow$.

Assume now that $b\mathcal{F}^\downarrow \subset b\mathcal{G}^\downarrow$. Using 4.11.15, proceed successively as follows:

$$\begin{aligned} b\mathcal{F}^\downarrow \subset b\mathcal{G}^\downarrow &\rightarrow (b\mathcal{F}^\downarrow)^\uparrow \subset (b\mathcal{G}^\downarrow)^\uparrow \\ &\rightarrow b(b\mathcal{F}^\downarrow)^\uparrow \subset b(b\mathcal{G}^\downarrow)^\uparrow \rightarrow b(b\mathcal{F}^\downarrow)^\uparrow \subset b(b\mathcal{G}^\downarrow)^\uparrow \rightarrow b\mathcal{F} \subset b\mathcal{G}. \end{aligned}$$

Finally, $\llbracket \mathcal{F} \subset \mathcal{G} \rrbracket \geq b \leftrightarrow b\mathcal{F}^\downarrow \subset b\mathcal{G}^\downarrow$, which completes the proof. \triangleright

4.11.17. Theorem. *Let $(\mathcal{F}_\xi)_{\xi \in \Xi}$ be a standard family of extensional filters and let $(b_\xi)_{\xi \in \Xi}$ be a standard partition of unity. A filter \mathcal{F} is the mixing of $(\mathcal{F}_\xi)_{\xi \in \Xi}$ by $(b_\xi)_{\xi \in \Xi}$ if and only if*

$$(\forall^{\text{st}} \xi \in \Xi) b_\xi \mu(\mathcal{F}) = b_\xi \mu(\mathcal{F}_\xi).$$

\triangleleft By definition, F belongs to the mixing $\sum_{\xi \in \Xi} b_\xi \mathcal{F}_\xi$ provided that there is a family $(F_\xi)_{\xi \in \Xi}$ such that $F_\xi \in \mathcal{F}_\xi$ ($\xi \in \Xi$) and, moreover, $F \supset \sum_{\xi \in \Xi} b_\xi F_\xi$. Since the members of the family $(\mathcal{F}_\xi)_{\xi \in \Xi}$ are extensional, by 4.8.1 and 4.8.2 conclude that \mathcal{F} is extensional too and the descent \mathcal{F}^\uparrow is the mixing of $(\mathcal{F}_\xi^\uparrow)_{\xi \in \Xi}$ by the same weights. Since $\mathbb{V}^{(B)}$ is a separated universe, we use 4.11.16 and 4.1.6 (5) to derive

$$\begin{aligned} \mathcal{F}^\uparrow &= \sum_{\xi \in \Xi} b_\xi \mathcal{F}_\xi^\uparrow \\ &\leftrightarrow (\forall^{\text{st}} \xi \in \Xi) b_\xi \mathcal{F}^\uparrow = b_\xi \mathcal{F}_\xi^\uparrow \\ &\leftrightarrow (\forall^{\text{st}} \xi \in \Xi) b_\xi \mathcal{F}^{\uparrow\downarrow} = b_\xi \mathcal{F}_\xi^{\uparrow\downarrow} \\ &\leftrightarrow (\forall^{\text{st}} \xi \in \Xi) b_\xi \mathcal{F} = b_\xi \mathcal{F}_\xi \leftrightarrow (\forall^{\text{st}} \xi \in \Xi) \mu(b_\xi \mathcal{F}) = \mu(b_\xi \mathcal{F}_\xi) \\ &\leftrightarrow (\forall^{\text{st}} \xi \in \Xi) b_\xi \mu(\mathcal{F}) = b_\xi \mu(\mathcal{F}_\xi). \end{aligned}$$

The proof is complete. \triangleright

Chapter 5

Infinitesimals and Subdifferentials

Infinitesimal analysis finds various applications in many areas of mathematics. In this chapter we discuss infinitesimals in *subdifferential calculus*, a branch of functional analysis which stems from the theory of extremal problems.

Convexifying plays a key role in optimization theory. The point is that we have the versatile and handy tools of convex analysis which have demonstrated their power and efficiency in the theoretical analysis and numerical solution of convex programs.

Local approximations to arbitrary sets and functions are the topic of non-smooth analysis which brings about a plenty of useful but complicated and often cumbersome formulas. The meaning of the new notions such as hypertangents, Rockafellar limits, and Clarke derivatives is difficult to comprehend from their formal definitions.

Infinitesimal analysis offers profound and effective simplifications by “*killing quantifiers*” which makes the perception of standard constructions easy and pleasant. We try to demonstrate this with providing an infinitesimal classification for one-sided tangents to arbitrary sets and functions. It is worth emphasizing that many constructions of this chapter have a wider range of applicability than subdifferential calculus and nonsmooth analysis.

5.1. Vector Topology

Studying local approximations involves vector topologies whose monads are the topic of this section. Unless otherwise stated, we assume the standard environment in the sequel.

5.1.1. *Let U be a standard star-like set in a vector space, i.e., $[0, 1] U \subset U$. The set U absorbs a standard set V if and only if $\alpha V \subset U$ for some (and hence all) positive infinitesimal α .*

◁ Since U absorbs V , there is some $\beta > 0$ satisfying $\beta V \subset U$ by definition. Since U and V are both standard, by transfer we conclude that $(\exists^{\text{st}}\beta > 0) \beta V \subset U$. If $\alpha > 0$ and $\alpha \approx 0$ then $\alpha V = \alpha/\beta(\beta V) \subset \alpha/\beta U \subset U$. The remaining claim is obvious. ▷

5.1.2. Let x be a standard element of a standard vector space X . The external set $\{\alpha x : \alpha > 0, \alpha \approx 0\}$ is the *conatus* or *direction* of x in X . The term “conatus” was minted by Hobbes [178, p. 173] who wrote that conatus “is motion through a space and a time less than any given, that is, less than any determined whether by exposition or assigned by number, that is, through a point.” The collection of all conatus of standard vectors of X is the *conatus* of X ; in symbols, $\text{cnt}(X)$.

5.1.3. A standard star-like set U is absorbing in X if and only if U includes the conatus $\text{cnt}(X)$ of X .

5.1.4. Let X be a standard vector space over the basic field \mathbb{F} , and let \mathcal{N} be a standard filter on X . There is a vector topology τ on X satisfying $\mathcal{N} = \tau(0)$ if and only if the monad $\mu(\mathcal{N})$ of \mathcal{N} includes the conatus $\text{cnt}(X)$ and, moreover, $\mu(\mathcal{N})$ is an external $\approx\mathbb{F}$ -submodule of X .

Here, as usual, $\approx\mathbb{F} := \{t \in \mathbb{F} : (\exists^{\text{st}}n \in \mathbb{N})|t| \leq n\}$ is the limited part of the basic field \mathbb{F} endowed with the natural structure of an external ring. Recall that \mathbb{F} is, as usual, either \mathbb{C} , the complex field or \mathbb{R} , the reals.

◁ \rightarrow : Since addition is continuous at zero, $\mu(\mathcal{N}) + \mu(\mathcal{N}) = \mu(\mathcal{N})$; i.e., $\mu(\mathcal{N})$ is an external subgroup of X . Take $\alpha \in \approx\mathbb{F}$ and let \mathcal{G} be a base for \mathcal{N} consisting of balanced sets. If $n \in {}^\circ\mathbb{N}$ satisfies $|\alpha| \leq n$ then $\alpha/n x \in G$ for $G \in {}^\circ\mathcal{G}$ and $x \in \mu(\mathcal{N})$. Therefore, $\alpha/n x \in \bigcap\{G : G \in {}^\circ\mathcal{G}\} = \mu(\mathcal{G}) = \mu(\mathcal{N})$; hence, $\alpha x \in n\mu(\mathcal{N}) = \mu(\mathcal{N})$; and, finally, $\alpha\mu(\mathcal{N}) = \mu(\mathcal{N})$ for $\alpha \in \approx\mathbb{F}$. Since \mathcal{N} has a base of absorbing sets; therefore, $\mu(\mathcal{N}) \supset \text{cnt}(X)$ by 5.1.3.

\leftarrow : Take $U \in {}^\circ\mathcal{N}$. By 4.1.4, this means that $U \supset \mu(\mathcal{N})$. If W is an infinitesimal member of \mathcal{N} then the balanced hull V of W is infinitesimal too (since $V \subset \mu(\mathcal{N})$). Moreover, $V + V \subset \mu(\mathcal{N}) + \mu(\mathcal{N}) \subset \mu(\mathcal{N}) \subset U$. Hence,

$$(\forall^{\text{st}}U \in \mathcal{N})(\exists V \in \mathcal{N})(V \text{ is balanced} \wedge V + V \subset U).$$

By transfer, conclude that $\mathcal{N} + \mathcal{N} = \mathcal{N}$ and, moreover, \mathcal{N} has a base of balanced sets. By 5.1.3, observe also that \mathcal{N} consists of standard balanced sets. Therefore, \mathcal{N} determines a vector topology on X . ▷

5.1.5. To each point x of the monad $\mu(X) := \mu(\tau(0))$ of a topological vector space X there is an unlimited natural number N satisfying $Nx \in \mu(X)$.

◁ If V is a standard neighborhood of zero and $n \in {}^\circ\mathbb{N}$, then (see 5.1.4) the set $A(n, V) := \{m \in \mathbb{N} : m \geq n \wedge mx \in V\}$ is nonempty (since $\mu(X) \subset V$). By transfer, there is an element N satisfying $(\forall^{\text{st}}n \in \mathbb{N})(\forall^{\text{st}}U \in \tau(0))(N \in A(n, V))$. This N is obviously a sought element. ▷

5.1.6. It is sometimes convenient considering *nearvector topologies* in applications. Such a topology τ on X is characterized by the following properties: first, multiplication of the vectors of X by every scalar of the basic field is continuous and, second, addition is also jointly continuous. If τ is a nearvector topology on X then the couple (X, τ) , as well as X itself, is a *neartopological vector space*. This notion is natural as is clear from the following easy proposition.

5.1.7. *Let X be a vector space over \mathbb{F} and let \mathcal{N} be a filter on X . There is a nearvector topology τ on X such that $\tau(0)$ coincides with \mathcal{N} if and only if the monad $\mu(\mathcal{N})$ of \mathcal{N} is an external vector space over the external field of standard scalars ${}^\circ\mathbb{F}$.*

5.1.8. In view of 5.1.7 we note that the monad of the neighborhood filter of the origin of a nearvector space is an external convex set. Every internal convex set U obviously contains arbitrary convex combinations of its elements: Given arbitrary finitely positive scalars $\alpha_1, \dots, \alpha_N$ with sum unity and a set $\{u_1, \dots, u_N\}$ of elements of U we have $\sum_{k=1}^N \alpha_k u_k \in U$. Here N is an arbitrary internal element of \mathbb{N} . This property is referred to as *hyperconvexity*. Observe that an external convex set may fail to be hyperconvex since it is impossible to use induction on internal naturals in the external universe. The corroborating examples are easy from the following useful proposition.

5.1.9. *A vector topology is locally convex if and only if the monad of its filter of zero neighborhoods is hyperconvex.*

$\langle \rightarrow$: Each standard neighborhood of a locally convex topology contains a standard convex and, hence, hyperconvex neighborhood. The intersection of hyperconvex external sets is hyperconvex too.

\leftarrow : Every standard neighborhood of zero contains the convex hull of every infinitesimal neighborhood of zero (since this hull is included in the monad of the neighborhood filter of zero and this monad is hyperconvex by hypothesis). Conclude by transfer that each neighborhood of zero contains a convex neighborhood of zero. \triangleright

5.1.10. Closing this section, we deviate slightly from the mainstream of exposition to observe that infinitesimal analysis of topological vector spaces and operators between them rests on studying the interlocation of various types of point. This relates to nearstandard and prenearstandard points as well as to new notions of a “bornological nature.” We list a few relevant facts.

5.1.11. *Let (X, τ) be a locally convex space, and let x be an internal point of X . The following are equivalent:*

- (1) $\alpha x \approx {}_\tau 0$ for every infinitesimal $\alpha \in \mathbb{F}$;

$$(2) \quad x \in \bigcap_{V \in \tau(0)} \bigcup_{n \in {}^\circ\mathbb{N}} nV;$$

(3) $p(x) \in {}^\approx\mathbb{F}$ for every standard continuous seminorm p (a member of the spectrum of τ).

◁ (1) \leftrightarrow (2): Proceed by the Nelson algorithm:

$$\begin{aligned} & (\forall \alpha \in \mathbb{F})(\alpha \approx 0 \rightarrow \alpha x \approx 0) \\ \leftrightarrow & (\forall^{\text{st}} V \in \tau(0))(\forall \alpha)(\forall^{\text{st}} n \in \mathbb{N})(|\alpha| \leq n^{-1} \rightarrow \alpha x \in V) \\ \leftrightarrow & (\forall^{\text{st}} V \in \tau(0))(\forall \alpha)(\exists^{\text{st}} n \in \mathbb{N})(|\alpha| \leq n^{-1} \rightarrow \alpha x \in V) \\ \leftrightarrow & (\forall^{\text{st}} V \in \tau(0))(\exists^{\text{st}} n \in \mathbb{N})(x \in nV). \end{aligned}$$

(1) \rightarrow (3): If p is a continuous seminorm then $|t|p(x) = p(|t|x) \approx 0$ for all $t \in {}^\approx\mathbb{R}$ by 4.2.7. Hence, $p(x) \in {}^\approx\mathbb{R}$.

(3) \rightarrow (1): Given a standard continuous seminorm p we have $p(\alpha x) = |\alpha|p(x) \approx 0$ whenever $|\alpha| \approx 0$. This means that αx is infinitesimal in the topology τ , which completes the proof. ▷

5.1.12. A point x satisfying one and hence all of the equivalent conditions 5.1.11 (1)–(3) is *limited* or *finite* in (X, τ) . In this event we write $x \in \text{ltd}(X, \tau)$ or simply $x \in \text{ltd}(X)$ whenever omitting the topology would entail no confusion. The external set $\text{ltd}(X)$ is the *limited* or *finite part* of X . More exact terms like “ τ -limited point” are also in common parlance.

5.1.13. Let X be a standard locally convex space. A standard set U in X is bounded if and only if U consists of limited points of X ; i.e., $U \subset \text{ltd}(X)$.

◁ \rightarrow : If U is bounded and p is a continuous seminorm on X then there is a standard $t \in {}^\circ\mathbb{R}$ satisfying $p(U) \leq t$. Given $\alpha \approx 0$ and $x \in U$, we thus have $p(\alpha x) \leq t\alpha$, i.e., $\alpha x \approx 0$.

\leftarrow : We now use the sequential test for a set to be bounded for the sake of diversity.

Let (α_n) be a standard vanishing sequence of scalars, and let (u_n) be a standard sequence of points of U . Check that $\alpha_n u_n \rightarrow 0$. To this end, let N be an unlimited natural. Then $\alpha_N \approx 0$ and so $\alpha_N u_N \approx 0$ by hypothesis and 5.1.11 (1). ▷

5.1.14. A point x of a space X is *bounded*, in symbols $x \in \text{bd}(X)$, provided that there is a standard bounded set containing x .

5.1.15. Let X be a separated locally convex space. The following are equivalent:

- (1) X is normable;
- (2) $\text{bd}(X) = \text{ltd}(X)$;
- (3) $\mu(X) \subset \text{bd}(X)$.

\triangleleft (1) \rightarrow (2): Obviously, $\text{bd}(X) \subset \text{ ltd}(X)$ with no hypothesis on X . If X is normable then $\text{ ltd}(X) = \{x \in X : \|x\| \in \approx\mathbb{R}\}$, where $\|\cdot\|$ is a suitable norm. Therefore, $\text{ ltd}(X)$ lies for instance in the unit ball $B_X := \{x \in X : \|x\| \leq 1\}$.

(2) \rightarrow (3): Since $\mu(X)$ is always a subset of $\text{ ltd}(X)$, the claim is obvious.

(3) \rightarrow (1): Let U be an infinitesimal neighborhood of zero in X . By hypothesis, to each $x \in U$ there is a standard set V such that V is bounded and $x \in V$. By idealization U lies in some bounded set. We are done on appealing to the celebrated Kolmogorov normability test. \triangleright

5.1.16. The above proposition shows in particular that if X is a general not necessarily normable space then X has more limited points than those bounded whereas $\text{ ltd}(X) = \text{ bd}(X)$ in a normable space X as follows from 5.1.15.

5.2. Classical Approximating and Regularizing Cones

Nonsmooth analysis intensively seeks convenient methods of local one-sided approximation to arbitrary functions and sets. The starting point of this search is Clarke's [263] definition of a subdifferential for a Lipschitz function.

The tangent cones and corresponding derivatives under study in nonsmooth analysis are often defined by cumbersome and bulky formulas. In this section we apply infinitesimal analysis as a method of "killing quantifiers," i.e. diminishing the complexity of formulas. Routinely assuming the standard environment (see 4.1.9–4.6.5), we show that the Bouligand, Clarke, and Hadamard cones as well as relevant regularizing cones result in fact from explicit infinitesimal constructions appealing straightforward to infinitely close points and directions.

5.2.1. Let X be a real vector space. Specifying some fixed nearvector topology $\sigma := \sigma_X$ in X with $\mathcal{N}_\sigma := \sigma(0)$ the neighborhood filter of zero, we also distinguish a nearvector topology τ with $\mathcal{N}_\tau := \tau(0)$. As usual, we introduce the infinite proximity on X that stems from the uniformity on X associated with σ by putting $x_1 \approx_\sigma x_2 \leftrightarrow x_1 - x_2 \in \mu(\mathcal{N}_\sigma)$. We do the same with τ . Moreover, we will assume that σ is a vector topology. For the monad of the neighborhood filter $\sigma(x)$ we use the symbol $\mu(\sigma(x))$ while denoting the monad $\mu(\sigma(0))$ simpler by $\mu(\sigma)$.

5.2.2. Given a subset F of X and a point x' in X , subdifferential calculus deals with the following *Hadamard*, *Clarke*, and *Bouligand cones*:

$$\text{Ha}(F, x') := \bigcup_{\substack{U \in \sigma(x') \\ \alpha' \\ \alpha}} \text{int}_\tau \left(\bigcap_{\substack{x \in F \cap U \\ 0 < \alpha \leq \alpha'}} \frac{F - x}{\alpha} \right);$$

$$\text{Cl}(F, x') := \bigcap_{V \in \mathcal{N}_\tau} \bigcup_{\substack{U \in \sigma(x') \\ \alpha' \\ \alpha}} \bigcap_{\substack{x \in F \cap U \\ 0 < \alpha \leq \alpha'}} \left(\frac{F - x}{\alpha} + V \right);$$

$$\text{Bo}(F, x') := \bigcap_{\substack{U \in \sigma(x') \\ \alpha'}} \text{cl}_\tau \left(\bigcup_{\substack{x \in F \cap U \\ 0 < \alpha \leq \alpha'}} \frac{F - x}{\alpha} \right),$$

where, as usual, $\sigma(x') := x' + \mathcal{N}_\sigma$. If $h \in \text{Ha}(F, x')$ then we sometimes say that F is *epi-Lipschitz* at x' in the direction of h .

Obviously,

$$\text{Ha}(F, x') \subset \text{Cl}(F, x') \subset \text{Bo}(F, x').$$

5.2.3. We also distinguish the *hypertangent cone*, the *cone of feasible directions* and the *contingency* of F at x' by the following relations:

$$\begin{aligned} \text{H}(F, x') &:= \bigcup_{\substack{U \in \sigma(x') \\ \alpha'}} \bigcap_{\substack{x \in F \cap U \\ 0 < \alpha \leq \alpha'}} \frac{F - x}{\alpha}; \\ \text{Fd}(F, x') &:= \bigcap_{\alpha' > 0} \frac{F - x'}{\alpha'}; \\ \text{K}(F, x') &:= \bigcap_{\alpha'} \text{cl}_\tau \left(\bigcup_{0 < \alpha \leq \alpha'} \frac{F - x'}{\alpha} \right). \end{aligned}$$

It stands to reason to assume $x' \in F$ for the sake of brevity. We then may say that $\text{H}(F, x')$ and $\text{K}(F, x')$ are the Hadamard and Bouligand cones when τ or σ is the discrete topology respectively. In the sequel we thus assume that $x' \in F$ and use the following abbreviations to save space:

$$\begin{aligned} (\forall^\bullet x) \varphi &:= (\forall x \approx_\sigma x') \varphi := (\forall x)(x \in F \wedge x \approx_\sigma x') \rightarrow \varphi, \\ (\forall^\bullet h) \varphi &:= (\forall h \approx_\tau h') \varphi := (\forall h)(h \in X \wedge h \approx_\tau h') \rightarrow \varphi, \\ (\forall^\bullet \alpha) \varphi &:= (\forall \alpha \approx 0) \varphi := (\forall \alpha)((\alpha > 0 \wedge \alpha \approx 0) \rightarrow \varphi). \end{aligned}$$

The quantifiers $\exists^\bullet x$, $\exists^\bullet h$, and $\exists^\bullet \alpha$ are defined by natural duality as follows

$$\begin{aligned} (\exists^\bullet x) \varphi &:= (\exists x \approx_\sigma x') \varphi := (\exists x)(x \in F \wedge x \approx_\sigma x') \wedge \varphi, \\ (\exists^\bullet h) \varphi &:= (\exists h \approx_\tau h') \varphi := (\exists h)(h \in X \wedge h \approx_\tau h') \wedge \varphi, \\ (\exists^\bullet \alpha) \varphi &:= (\exists \alpha \approx 0) \varphi := (\exists \alpha)(\alpha > 0 \wedge \alpha \approx 0) \wedge \varphi. \end{aligned}$$

We now establish that the above cones admit simple descriptions on using infinitesimals.

5.2.4. *The Bouligand cone is the standardization of the $\exists\exists\exists$ -cone; i.e.,*

$$h' \in \text{Bo}(F, x') \leftrightarrow (\exists^\bullet x)(\exists^\bullet \alpha)(\exists^\bullet h)(x + \alpha h \in F)$$

for a standard point h' .

◁ From the definition of the Bouligand cone it follows that

$$\begin{aligned} h' \in \text{Bo}(F, x') & \\ \leftrightarrow (\forall U \in \sigma(x'))(\forall \alpha' \in \mathbb{R})(\forall V \in \mathcal{N}_\tau)(\exists x \in F \cap U) & \\ (\exists 0 < \alpha \leq \alpha')(\exists h \in h' + V)(x + \alpha h \in F) & \\ \leftrightarrow (\forall U)(\forall \alpha')(\forall V)(\exists x)(\exists \alpha)(\exists h) & \\ (x \in F \cap U \wedge h \in h' + V \wedge 0 < \alpha \leq \alpha' \wedge x + \alpha h \in F). & \end{aligned}$$

By transfer,

$$\begin{aligned} h' \in \text{Bo}(F, x') \leftrightarrow (\forall^{\text{st}} U)(\forall^{\text{st}} \alpha')(\forall^{\text{st}} V)(\exists^{\text{st}} x)(\exists^{\text{st}} \alpha)(\exists^{\text{st}} h) & \\ (x \in F \cap U \wedge h \in h' + V \wedge 0 < \alpha \leq \alpha' \wedge x + \alpha h \in F). & \end{aligned}$$

By idealization,

$$\begin{aligned} h' \in \text{Bo}(F, x') \rightarrow (\exists x)(\exists \alpha)(\exists h)(\forall^{\text{st}} U)(\forall^{\text{st}} \alpha')(\forall^{\text{st}} V) & \\ (x \in F \cap U \wedge h \in h' + V \wedge 0 < \alpha \leq \alpha' \wedge x + \alpha h \in F) & \\ \rightarrow (\exists x \approx_\sigma x')(\exists \alpha \approx 0)(\exists h \approx_\tau h')(x + \alpha h \in F) & \\ \rightarrow (\exists^\bullet x)(\exists^\bullet \alpha)(\exists^\bullet h)(x + \alpha h \in F). & \end{aligned}$$

Assume conversely that a standard element h' belongs to the standardization of the “ $\exists\exists\exists$ -cone.” Since every standard element of a standard filter contains the monad of this filter; therefore,

$$\begin{aligned} (\forall^{\text{st}} U \in \sigma(x'))(\forall^{\text{st}} \alpha' \in \mathbb{R})(\forall^{\text{st}} V \in \mathcal{N}_\tau) & \\ (\exists x \in F \cap U)(\exists 0 < \alpha < \alpha')(\exists h \in h' + V)(x + \alpha h \in F). & \end{aligned}$$

By transfer, $h' \in \text{Bo}(F, x')$, which completes the proof. ▷

5.2.5. Proposition 5.2.4 may be rewritten as

$$\text{Bo}(F, x') = * \{h' \in X : (\exists^\bullet x)(\exists^\bullet \alpha)(\exists^\bullet h)(x + \alpha h \in F)\}$$

with $*$ standing for standardization. This leads to a more impressive notation:

$$\exists\exists\exists(F, x') := \text{Bo}(F, x').$$

We will proceed likewise in the sequel without circumlocution.

5.2.6. *The Hadamard cone is the standardization of the $\forall\forall\forall$ -cone:*

$$\text{Ha}(F, x') = \forall\forall\forall(F, x').$$

In other words,

$$h' \in \text{Ha}(F, x') \leftrightarrow (x' + \mu(\sigma)) \cap F + \mu(\mathbb{R}_+)(h' + \mu(\tau)) \subset F,$$

whenever h' , F , and x' are standard, with $\mu(\mathbb{R}_+)$ the monad comprising positive infinitesimals in \mathbb{R} .

◁ The claim ensues from 5.2.4 by duality on forgetting the presence of F in $\exists^\bullet x$, which is by all means legitimate. ▷

5.2.7. The above implies also that

$$\begin{aligned} h' \in \text{Ha}(F, x') &\leftrightarrow (\forall^\bullet x)(\forall^\bullet \alpha)(x + \alpha h' \in F), \\ h' \in \text{K}(F, x') &\leftrightarrow (\exists^\bullet \alpha)(\exists^\bullet h)(x' + \alpha h \in F). \end{aligned}$$

5.2.8. *If h' , F , and x' are standard then, assuming weak idealization, the following are equivalent:*

- (1) $h' \in \text{Cl}(F, x')$;
- (2) there are infinitely small $U \in \sigma(x')$, $V \in \mathcal{N}_\tau$, and $\alpha' > 0$ satisfying

$$h' \in \bigcap_{\substack{0 < \alpha \leq \alpha' \\ x \in F \cap U}} \left(\frac{F - x}{\alpha} + V \right);$$

- (3) $(\exists U \in \sigma(x'))(\exists \alpha')(\forall x \in F \cap U)(\forall 0 < \alpha \leq \alpha')(\exists h \approx_\tau h')(x + \alpha h \in F)$.

◁ Deciphering the definition and using obvious abbreviations, note that

$$\begin{aligned} &h' \in \text{Cl}(F, x') \\ \leftrightarrow &(\exists V)(\exists U)(\exists \alpha')(\forall x \in F \cap U)(\forall 0 < \alpha \leq \alpha')(\exists h \in h' + V) \\ &(x + \alpha h \in F). \end{aligned}$$

By transfer and weak idealization, infer

$$\begin{aligned} h \in \text{Cl}(F, x') &\rightarrow (\forall^{\text{st}} V)(\exists^{\text{st}} U)(\exists^{\text{st}} \alpha')(\forall x \in F \cap V) \\ &(\forall 0 < \alpha \leq \alpha')(\exists h \in h' + V)(x + \alpha h \in F) \\ &\rightarrow (\forall^{\text{st}} \{V_1, \dots, V_n\})(\exists^{\text{st}} U)(\exists^{\text{st}} \alpha')(\exists^{\text{st}} V)(\forall k := 1, \dots, n) \end{aligned}$$

$$\begin{aligned}
V_k \supset V \wedge (\forall x \in F \cap U)(\forall 0 < \alpha \leq \alpha')(\exists h \in h' + V)(x + \alpha h \in F) \\
\rightarrow (\exists U)(\exists \alpha')(\exists V)(\forall^{\text{st}} V') V' \supset V \wedge (\forall x \in F \cap U) \\
(\forall 0 < \alpha \leq \alpha')(\exists h \in h' + V)(x + \alpha h \in F).
\end{aligned}$$

From (1) it is now easy that there are an infinitesimal α and some $V \in \mathcal{N}_\tau$, $V \subset \mu(\tau)$ $U \in \sigma(x')$, $U \subset \mu(\sigma) + x'$ meeting (2) and hence (3).

Conversely, if (3) holds then from the definition of \approx_τ obtain

$$\begin{aligned}
(\forall^{\text{st}} V)(\exists U)(\exists \alpha')(\forall x \in F \cap U)(\forall 0 < \alpha \leq \alpha')(\exists h \in h' + V) \\
(x + \alpha h \in F).
\end{aligned}$$

By transfer, $h' \in \text{Cl}(F, x')$, which ends the proof. \triangleright

5.2.9. *The Clarke cone is the $\forall\forall\exists$ -cone*

$$\text{Cl}(F, x') = \forall\forall\exists(F, x')$$

under the assumption of strong idealization. In other words,

$$h' \in \text{Cl}(F, x') \leftrightarrow (\forall^\bullet x)(\forall^\bullet \alpha)(\exists^\bullet h)(x + \alpha h \in F).$$

\triangleleft Take $h' \in \text{Cl}(F, x')$. Choose $x \approx_\sigma x'$ and $\alpha > 0$, $\alpha \approx 0$ arbitrarily. To each standard member V of the neighborhood filter \mathcal{N}_τ , by transfer there is some h satisfying $h \in h' + V$ and $x + \alpha h \in F$. By idealization,

$$\begin{aligned}
(\forall^{\text{st}} V)(\exists h)(h \in h' + V \wedge x + \alpha h \in F) \\
\rightarrow (\exists h)(\forall^{\text{st}} V)(h \in h' + V \wedge x + \alpha h \in F) \rightarrow (\exists^\bullet h)(x + \alpha h \in F),
\end{aligned}$$

i.e., $h' \in \forall\forall\exists(F, x')$.

Given $h' \in \forall\forall\exists(F, x')$, take an arbitrary standard member V of the neighborhood filter \mathcal{N}_τ , some infinitesimal neighborhood U of x' , and a positive infinitesimal α' . By hypothesis,

$$(\exists x \in F \cap U)(\forall 0 < \alpha \leq \alpha')(x + \alpha h \in F)$$

for some $h \approx_\tau h'$. In other words,

$$(\forall^{\text{st}} V)(\exists U)(\exists \alpha')(\forall x \in F \cap U)(\forall 0 < \alpha \leq \alpha')(\exists h \in h' + V)(x + \alpha h \in F).$$

By transfer, conclude that $h' \in \text{Cl}(F, x')$. \triangleright

5.2.10. We now apply the above “infinitesimal” test for the members of the Clarke cone to deducing its basic (and well-known) property. A more general statement will appear below.

5.2.11. *The Clarke cone at each point of an arbitrary set in a topological vector space is convex and closed.*

◁ By transfer, we may assume the standard environment in which all parameters (space, topology, set, etc.) are standard.

To prove closure, take $h_0 \in \text{cl}_\tau \text{Cl}(F, x')$. Given a standard member V of the neighborhood filter \mathcal{N}_τ and standard elements $V_1, V_2 \in \mathcal{N}_\tau$ satisfying $V_1 + V_2 \subset V$, find a standard element $h' \in \text{Cl}(F, x')$ such that $h' - h_0 \in V'$. Moreover, for all $x \approx_\tau x'$ and $\alpha > 0$, $\alpha \approx 0$ there is some h such that $h \in h' + V_2$ and $x + \alpha h \in F$. Obviously, $h \in h' + V_2 \subset h_0 + V_1 + V_2 \subset h_0 + V$ and so $h_0 \in \text{Cl}(F, x')$.

To prove convexity it suffices to note that $\mu(\tau) + \mu(\mathbb{R}_+)\mu(\tau) \subset \mu(\tau)$, since $(x, \alpha, h) \mapsto x + \alpha h$ is a continuous mapping. ▷

5.2.12. *Assume that θ is a vector topology and $\theta \geq \tau$. Then*

$$\forall \forall \exists (\text{cl}_\theta F, x') \subset \forall \forall \exists (F, x').$$

Moreover, if $\theta \geq \sigma$ then

$$\forall \forall \exists (\text{cl}_\theta F, x') = \forall \forall \exists (F, x').$$

◁ Let $h' \in \forall \forall \exists (\text{cl}_\theta F, x')$ be a standard element of the cone in question. Choose elements $x \in F$ and $\alpha > 0$ such that $x \approx_\sigma x'$ and $\alpha \approx 0$. Clearly, $x \in \text{cl}_\theta F$. Hence, $x + \alpha h \in \text{cl}_\theta F$ for some $h \approx_\tau h'$. Consider an infinitely small neighborhood W included in $\mu(\theta)$. The neighborhood αW is also an element of $\theta(0)$. Thus, $x'' - (x + \alpha h) \in \alpha W$ for some $x'' \in F$. Put $h'' := (x'' - x)/\alpha$. Obviously, $x + \alpha h'' \in F$ and $\alpha h'' \in \alpha h + \alpha W$. Therefore,

$$h'' \in h + W \subset h' + \mu(\tau) + W \subset h' + \mu(\tau) + \mu(\theta) \subset h' + \mu(\tau) + \mu(\tau) \subset h' + \mu(\tau);$$

i.e., $h'' \approx_\tau h'$. Hence, $h' \in \forall \forall \exists (F, x')$.

Assume now that $\theta \geq \sigma$ and $h' \in \forall \forall \exists (F, x')$. Choose an arbitrary infinitesimal α and an element $x \in \text{cl}_\theta F$ such that $x \approx_\sigma x'$. Find $x'' \in F$ satisfying $x - x'' \in \alpha W$, with $W \subset \mu(\theta)$ an infinitesimal symmetric neighborhood of the origin in θ . Since $\theta \geq \sigma$; therefore, $\mu(\theta) \subset \mu(\sigma)$, i.e. $x - x'' \in \mu(\theta) \subset \mu(\sigma)$ or, in other words, $x \approx_\sigma x' \approx_\sigma x''$. Considering h' standard, note that $x'' + \alpha h \in F$ for some $h \approx_\sigma h'$ by definition. Putting $h'' := (x'' - x)/\alpha + h$, infer

$$h'' \in h + W \subset h + \mu(\theta) \subset h' + \mu(\theta) + \mu(\tau) \subset h' + \mu(\tau) + \mu(\tau) \subset h' + \mu(\tau),$$

i.e., $h'' \approx_\tau h'$. Moreover,

$$x + \alpha h'' = x + (x'' - x) + \alpha h = x'' + \alpha h \in \text{cl}_\theta F.$$

Finally, $h' \in \forall \forall \exists (\text{cl}_\theta F, x')$. ▷

5.2.13. The above representation implies in particular that

$$\text{Ha}(F, x') \subset \text{H}(F, x') \subset \text{Cl}(F, x') \subset \text{K}(F, x') \subset \text{cl}_\tau \text{Fd}(F, x').$$

If $\sigma = \tau$ and F is a convex set then

$$\text{Fd}(F, x') \subset \text{Cl}(F, x') \subset \text{cl Fd}(F, x');$$

whence

$$\text{Cl}(F, x') = \text{K}(F, x') = \text{cl Fd}(F, x').$$

5.2.14. The nonstandard definitions for the Bouligand, Hadamard, and Clarke cones show that these cones belong to the list of eight possible cones with the infinitesimal prefix $(\mathbf{Q}^\bullet x)(\mathbf{Q}^\bullet \alpha)(\mathbf{Q}^\bullet h)$, with \mathbf{Q}^\bullet standing for \forall or \exists . To describe all these cones completely it suffices to characterize $\forall\exists\exists$ -cones and $\forall\exists\forall$ -cones.

5.2.15. *The following holds:*

$$\forall\exists\exists(F, x') = \bigcap_{\substack{\alpha' \\ \forall \in \mathcal{N}_\tau}} \bigcup_{U \in \sigma(x')} \bigcap_{x \in F \cap U} \left(V + \bigcup_{0 < \alpha \leq \alpha'} \frac{F - x}{\alpha} \right).$$

\triangleleft To prove, observe first that the sought equality abbreviates the following: if h' , F , and x' are standard then

$$\begin{aligned} & (\forall^\bullet x)(\exists^\bullet \alpha)(\exists^\bullet h)(x + \alpha h \in F) \\ \leftrightarrow & (\forall V \in \mathcal{N}_\tau)(\forall \alpha')(\exists U \in \sigma(x'))(\forall x \in F \cap U) \\ & (\exists 0 < \alpha \leq \alpha')(\exists h \in h' + V)(x + \alpha h \in F). \end{aligned}$$

Therefore, given $h' \in \forall\exists\exists(F, x')$, a standard V in \mathcal{N}_τ , and a standard $\alpha > 0$, we may take an internal subset of the monad $\mu(\sigma(x'))$ as the sought neighborhood U . By transfer and idealization,

$$\begin{aligned} & (\forall^{\text{st}} V)(\forall^{\text{st}} \alpha')(\forall x \approx_\sigma x')(\exists 0 < \alpha \leq \alpha')(\exists h \in h' + V)(x + \alpha h \in F) \\ & \rightarrow (\forall x \approx_\sigma x')(\forall^{\text{st}} \{V_1, \dots, V_n\})(\forall^{\text{st}} \{\alpha'_1, \dots, \alpha'_n\}) \\ & (\exists h)(\exists \alpha)(\forall k := 1, \dots, n)(0 < \alpha \leq \alpha'_k \wedge h \in h' + V_k \wedge x + \alpha h \in F) \\ & \rightarrow (\forall x \approx_\sigma x')(\exists h)(\exists \alpha)(\forall^{\text{st}} V)(h \in h' + V) \wedge (\forall^{\text{st}} \alpha') \\ & (0 < \alpha \leq \alpha' \wedge x + \alpha h \in F) \rightarrow (\forall^\bullet x)(\exists^\bullet h)(\exists \alpha \approx 0)(x + \alpha h \in F) \\ & \rightarrow h' \in {}^* \{h' : (\forall^\bullet x)(\exists^\bullet \alpha)(\exists^\bullet h)(x + \alpha h \in F)\} \rightarrow h' \in \forall\exists\exists(F, x'), \end{aligned}$$

which completes the proof. \triangleright

5.2.16. Alongside with the eight infinitesimal cones of the above classical series discussed above, there are nine more couples of cones containing the Hadamard cone and lying in the Bouligand cone. These cones result evidently from changing the order of quantifiers. Five of these couples are constructed in a somewhat bizarre manner by analogy with the $\forall\exists\forall$ -cone, the remaining couples are generated by permutations and dualizations of the Clarke cone and the $\forall\exists\exists$ -cone.

Using natural notation, we for instance infer

$$\begin{aligned}\forall\alpha\forall h\exists x(F, x') &= \bigcap_{U \in \sigma(x')} \bigcup_{\alpha'} \text{int}_\tau \left(\bigcap_{0 < \alpha \leq \alpha'} \bigcup_{x \in F \cap U} \frac{F - x}{\alpha} \right), \\ \exists h\exists x\forall\alpha(F, x') &= \bigcup_{\alpha'} \bigcap_{U \in \sigma(x')} \text{cl}_\tau \left(\bigcup_{x \in F \cap U} \bigcap_{0 < \alpha \leq \alpha'} \frac{F - x}{\alpha} \right), \\ \exists h\forall x\forall\alpha(F, x') &= \bigcap_{\substack{U \in \sigma(x') \\ \alpha'}} \text{cl}_\tau \left(\bigcap_{\substack{0 < \alpha \leq \alpha' \\ x \in F \cap U}} \frac{F - x}{\alpha} \right).\end{aligned}$$

The last cone, narrower than the Clarke cone, is convex provided that $\mu(\sigma) + \mu(\mathbb{R}_+)\mu(\tau) \subset \mu(\sigma)$. We denote this cone by $\text{Ha}^+(F, x')$. Observe that

$$\text{Ha}(F, x') \subset \text{Ha}^+(F, x') \subset \text{Cl}(F, x').$$

Also convex is the $\forall\alpha\exists h\forall x$ -cone, denoted by $\text{In}(F, x')$. Obviously,

$$\text{Ha}^+(F, x) \subset \text{In}(F, x') \subset \text{Cl}(F, x').$$

5.2.17. Calculation of tangents to the composite of correspondences rests on special *regularizing cones*.

Assume that $F \subset X \times Y$, where X and Y are vector spaces furnished with topologies σ_X, τ_X and σ_Y, τ_Y . Take $a' := (x', y') \in F$. Assigning $\sigma := \sigma_X \times \sigma_Y$, put

$$\begin{aligned}\mathbf{R}^1(F, a') &:= \bigcap_{V \in \mathcal{N}_{\tau_Y}} \bigcup_{\substack{W \in \sigma(a') \\ \alpha'}} \bigcap_{\substack{a \in W \cap F \\ 0 < \alpha \leq \alpha'}} \left(\frac{F - a}{\alpha} + 0 \times V \right), \\ \mathbf{Q}^1(F, a') &:= \bigcap_{V \in \mathcal{N}_{\tau_Y}} \bigcup_{\substack{W \in \sigma(a') \\ \alpha' \\ U \in \mathcal{N}_\sigma}} \bigcap_{\substack{a \in W \cap F \\ 0 < \alpha \leq \alpha' \\ x \in U}} \left(\frac{F - a}{\alpha} + \{x\} \times V \right),\end{aligned}$$

$$\text{QR}^2(F, a') := \bigcup_{\substack{W \in \sigma(a') \\ \alpha' \\ U \in \mathcal{N}_\sigma}} \bigcap_{\substack{a \in W \cap F \\ 0 < \alpha \leq \alpha' \\ x \in U}} \left(\frac{F - a}{\alpha} + (x, 0) \right).$$

We must confess here in one of our major sins: We tacitly identify 0 and the singleton $\{0\}$, whereas still discriminating between x and $\{x\}$. This popular abuse is very convenient and we exercise it throughout the book with no remorse.

The cones $\text{R}^2(F, a')$, $\text{Q}^2(F, a')$, and $\text{QR}^1(F, a')$ are defined by duality. Moreover, we use analogous notation for the case of the product of more than two spaces: The superindex of the symbol of an approximating set signifies the coordinate on which we impose the corresponding condition. Note also that applications usually involve pairwise coincident topologies: $\sigma_X = \tau_X$ and $\sigma_Y = \tau_Y$.

We proceed with nonstandard definitions of regularizing cones.

5.2.18. *If $s' \in X$ and $t' \in Y$ are standard then*

$$\begin{aligned} & (s', t') \in \text{R}^1(F, a') \\ \leftrightarrow & (\forall a \approx_\sigma a', a \in F)(\forall \alpha \in \mu(\mathbb{R}_+))(\exists t \approx_{\tau_Y} t')(a + \alpha(s', t) \in F); \\ & (s', t') \in \text{Q}^1(F, a') \\ \leftrightarrow & (\forall a \approx_\sigma a', a \in F)(\forall \alpha \in \mu(\mathbb{R}_+))(\forall s \approx_{\tau_X} s')(\exists t \approx_{\tau_Y} t')(a + \alpha(s, t) \in F); \\ & (s', t') \in \text{QR}^2(F, a') \\ \leftrightarrow & (\forall a \approx_\sigma a', a \in F)(\forall \alpha \in \mu(\mathbb{R}_+))(\forall s \approx_{\tau_X} s')(a + \alpha(s, t') \in F). \end{aligned}$$

5.2.19. As seen from 5.2.18, the cones of the type QR^j are variations of the Hadamard cone, while the cones R^j are particular cases of the Clarke cone. The cones R^j are also specializations of the cones of the type Q^j if we appropriately choose discrete topologies. These cones are convex under the routine assumptions. We demonstrate this claim only for the cone Q^j , which is clearly enough.

5.2.20. *If $(a, \alpha, b) \mapsto a + \alpha b$ is a continuous mapping from $(X \times Y, \sigma) \times (\mathbb{R}, \tau_{\mathbb{R}}) \times (X \times Y, \tau_X \times \tau_Y)$ to $(X \times Y, \sigma)$ then $\text{Q}^j(F, a')$ is a convex cone for $j := 1, 2$.*

◁ By transfer, we may proceed in the standard environment, and so the tests of 5.2.18 are readily available.

Take (s', t') and (s'', t'') in $\text{Q}^1(F, x')$. Put $a \approx_\sigma a'$, $a \in F$. Let $\alpha \approx 0$ be positive and $s \approx_{\tau_X} (s' + s'')$. By 5.2.18, $a_1 := a + \alpha(s - s'', t_1) \in F$ for some $t_1 \approx_{\tau_Y} t'$. By hypothesis, $\mu(\sigma) + \alpha(\mu(\tau_X) \times \mu(\tau_Y)) \subset \mu(\sigma)$. Therefore, $a_1 \approx_\sigma a$ and $a_1 \in F$. Applying 5.2.18 again, find $t_2 \approx_{\tau_Y} t''$ satisfying $a_1 + \alpha(s'', t_2) \in F$. Putting $t := t_1 + t_2$, observe that $t \approx_{\tau_Y} (t' + t'')$ and

$$a + \alpha(s, t) = a + \alpha(s - s'', t_1) + \alpha(s'', t_2) = a_1 + \alpha(s'', t_2) \in F,$$

which completes the proof since the homogeneity of $Q^1(F, a')$ ensues from closure of the monad of a nearvector topology under multiplication by standard scalars (cf. 5.1.4). \triangleright

5.2.21. The above analysis corroborates the introduction of the cones P^j and S^j by standardization as follows:

$$\begin{aligned} & (s', t') \in P^2(F, a') \\ \leftrightarrow & (\exists s \approx_{\tau_X} s')(\forall t \approx_{\tau_Y} t')(\forall a \approx_{\sigma} a', a \in F)(\forall \alpha \in \mu(\mathbb{R}_+))(a + \alpha(s, t) \in F), \\ (s', t') \in S^2(F, a) \leftrightarrow & (\forall t \approx_{\tau_Y} t')(\exists s \approx_{\tau_X} s')(\forall a \approx_{\sigma} a', a \in F)(\forall \alpha \in \mu(\mathbb{R}_+)) \\ & (a + \alpha(s, t) \in F). \end{aligned}$$

The explicit definitions of P^j and S^j are available in principle (we will discuss this in the subsection to follow). However, the arising formulas (especially that for S^j) are enormously cumbersome and bulky, and so, of little avail. Moreover, as we have seen already, these explicit formulas often obscure analysis by hiding the transparent “infinitesimal” ideas behind the formal constructions.

5.2.22. *The inclusions hold*

$$\text{Ha}(F, a') \subset P^j(F, a') \subset S^j(F, a') \subset Q^j(F, a') \subset R^j(F, a') \subset \text{Cl}(F, a')$$

for $j := 1, 2$. These cones are convex provided that $\mu(\sigma) + \alpha(\mu(\tau_X) \times \mu(\tau_Y)) \subset \mu(\sigma)$ for all $\alpha > 0$, $\alpha \approx 0$.

\triangleleft The inclusions are immediate from the nonstandard definitions of these cones.

As regards convexity, we have already demonstrated it for most of these cones. For the sake of perfection, we now show that $S^2(F, a')$ is a convex cone.

The set $S^2(F, a')$ is closed under multiplication by every positive standard real since the monad of real infinitesimals is indivisible. We are left with checking that $S^2(F, a')$ is a semigroup. To this end, given standard members (s', t') and (s'', t'') of $S^2(F, a')$, choose $t \approx_{\tau_Y} (t' + t'')$. Then $t - t'' \approx_{\tau_Y} t'$ and there is an $s_1 \approx_{\tau_X} s'$ serving to $t - t''$ by the definition of $S^2(F, a')$. Choose a point $s_2 \approx_{\tau_X} s''$ that “serves” to t'' by the same definition. Clearly, $(s_1 + s_2) \approx_{\tau_X} (s' + s'')$. In this event $a_1 := a + \alpha(s_1, t - t'') \in F$ for all $a \in F$ and $\alpha > 0$ such that $a \approx_{\sigma} a'$ and $\alpha \approx 0$. Since a_1 is infinitely close to a' with respect to σ , from the choice of s_2 we conclude that $a_1 + \alpha(s_2, t'') \in F$. Hence, $a + \alpha(s_1 + s_2, t) \in F$, i.e., $(s' + s'', t' + t'') \in S^2(F, a')$.

An analogous straightforward argument proves that $P^j(F, a')$ is a convex cone too. \triangleright

5.2.23. Inspection of the proof of 5.2.22 prompts us to consider convex “enlargements” of P^j and S^j , i.e. the cones P^{+j} and S^{+j} that result from transferring the quantifier $\forall\alpha$. For instance, the cone $P^{+2}(F, a')$ is defined as

$$(s', t') \in P^{+2}(F, a') \leftrightarrow (\forall \alpha \in \mu(\mathbb{R}_+))(\exists s \approx_{\tau_X} s')(\forall t \approx_{\tau_Y} t') \\ (\forall a \approx_{\sigma} a', a \in F)(a + \alpha(s, t) \in F).$$

In view of 5.2.19 it is reasonable to use the regularizations that result from specifying the cone Ha^+ by choosing discrete topologies. We skip the corresponding explicit formulas. The importance of regularizing cones stems from their role in subdifferentiation of composite mappings we address in 5.5.

5.3. Kuratowski and Rockafellar Limits

The preceding section shows that many vital constructions result from alternating quantifiers over infinitesimals. Similar effects arise in various problems and pertain to principal facts. We will now address those which are most often in subdifferential calculus. We start with some general observations concerning the Nelson algorithm.

5.3.1. Assume that $\varphi = \varphi(x, y) \in (\text{ZFC})$; i.e., φ is some formula of ZFC with no bound variables but x, y . Then

$$(\forall x \in \mu(\mathcal{F}))\varphi(x, y) \leftrightarrow (\exists^{\text{st}} F \in \mathcal{F})(\forall x \in F)\varphi(x, y), \\ (\exists x \in \mu(\mathcal{F}))\varphi(x, y) \leftrightarrow (\forall^{\text{st}} \in \mathcal{F})(\exists x \in F)\varphi(x, y),$$

with $\mu(\mathcal{F})$ standing as usual for the monad of a standard filter \mathcal{F} .

◁ It suffices to demonstrate the implication \rightarrow in the first of the equivalences.

By hypothesis, the internal property $\psi := (\forall x \in F)\varphi(x, y)$ is fulfilled for every remote element F of \mathcal{F} . Hence, ψ is valid for some standard F by the Cauchy principle. ▷

5.3.2. Assume that $\varphi = \varphi(x, y, z) \in \text{ZFC}$. Assume further that \mathcal{F} and \mathcal{G} are standard filters on some standard sets. In this case

$$(\forall x \in \mu(\mathcal{F}))(\exists y \in \mu(\mathcal{G}))\varphi(x, y, z) \\ \leftrightarrow (\forall^{\text{st}} G \in \mathcal{G})(\exists^{\text{st}} F \in \mathcal{F})(\forall x \in F)(\exists y \in G)\varphi(x, y, z) \\ \leftrightarrow (\exists^{\text{st}} F(\cdot))(\forall^{\text{st}} G \in \mathcal{G})(\forall x \in F(G))(\exists y \in G)\varphi(x, y, z), \\ (\exists x \in \mu(\mathcal{F}))(\forall y \in \mu(\mathcal{G}))\varphi(x, y, z) \\ \leftrightarrow (\exists^{\text{st}} G \in \mathcal{G})(\forall^{\text{st}} F \in \mathcal{F})(\exists x \in F)(\forall y \in G)\varphi(x, y, z) \\ \leftrightarrow (\forall^{\text{st}} F(\cdot))(\exists^{\text{st}} G \in \mathcal{G})(\exists x \in F(G))(\forall y \in G)\varphi(x, y, z),$$

with $F(\cdot)$ symbolizing a function from \mathcal{G} to \mathcal{F} .

◁ The proof consists in appealing to the idealization and construction principles together with 5.3.1. ▷

5.3.3. Take $\varphi = \varphi(x, y, z, u) \in (\text{ZFC})$ and let \mathcal{F}, \mathcal{G} , and \mathcal{H} be three standard filters. Assuming the standard environment, the following hold:

$$\begin{aligned}
 & (\forall x \in \mu(\mathcal{F}))(\exists y \in \mu(\mathcal{G}))(\forall z \in \mu(\mathcal{H})) \varphi(x, y, z, u) \\
 & \leftrightarrow (\forall G(\cdot))(\exists F \in \mathcal{F})(\exists^{\text{fin}} \mathcal{H}_0 \subset \mathcal{H})(\forall x \in \mathcal{F}) \\
 & \quad (\exists H \in \mathcal{H}_0)(\exists y \in G(H))(\forall z \in H) \varphi(x, y, z, u), \\
 & (\exists x \in \mu(\mathcal{F}))(\forall y \in \mu(\mathcal{G}))(\exists z \in \mu(\mathcal{H})) \varphi(x, y, z, u) \\
 & \leftrightarrow (\exists G(\cdot))(\forall F \in \mathcal{F})(\forall^{\text{fin}} \mathcal{H}_0 \subset \mathcal{H})(\exists x \in F) \\
 & \quad (\forall H \in \mathcal{H}_0)(\forall y \in G(H))(\exists z \in H) \varphi(x, y, z, u),
 \end{aligned}$$

with $G(\cdot)$ standing for a function from \mathcal{H} to \mathcal{G} .

◁ The Nelson algorithm yields

$$\begin{aligned}
 & (\forall x \in \mu(\mathcal{F}))(\exists y \in \mu(\mathcal{G}))(\forall z \in \mu(\mathcal{H})) \varphi \\
 \leftrightarrow & (\forall x \in \mu(\mathcal{F}))(\forall^{\text{st}} G(\cdot))(\exists^{\text{st}} H \in \mathcal{H})(\exists y \in G(H))(\forall z \in H) \varphi \\
 \leftrightarrow & (\forall^{\text{st}} G(\cdot))(\forall x) (\exists^{\text{st}} F \in \mathcal{F})(\exists^{\text{st}} H \in \mathcal{H}) \\
 & \quad (x \in F \rightarrow (\exists y \in G(H))(\forall z \in H) \varphi) \\
 \leftrightarrow & (\forall^{\text{st}} G(\cdot))(\exists^{\text{st fin}} \mathcal{F}_0)(\exists^{\text{st fin}} \mathcal{H}_0)(\forall x) (\exists F \in \mathcal{F}_0)(\exists H \in \mathcal{H}_0) \\
 & \quad (F \in \mathcal{F} \wedge H \in \mathcal{H} \wedge (x \in F \rightarrow (\exists y \in G(H))(\forall z \in H) \varphi)) \\
 \leftrightarrow & (\forall^{\text{st}} G(\cdot))(\exists^{\text{st fin}} \mathcal{F}_0 \subset \mathcal{F})(\exists^{\text{st fin}} \mathcal{H}_0 \subset \mathcal{H})(\forall x) (\exists F \in \mathcal{F}_0) \\
 & \quad (x \in F \rightarrow (\exists H \in \mathcal{H}_0)(\exists y \in G(H))(\forall z \in H) \varphi) \\
 \leftrightarrow & (\forall G(\cdot))(\exists^{\text{fin}} \mathcal{F}_0 \subset \mathcal{F})(\exists^{\text{fin}} \mathcal{H}_0 \subset \mathcal{H})(\forall x) \\
 \leftrightarrow & ((\forall F \in \mathcal{F}_0) x \in F \rightarrow (\exists H \in \mathcal{H}_0) (\exists y \in G(H))(\forall z \in H) \varphi) \\
 \leftrightarrow & (\forall G(\cdot))(\exists^{\text{fin}} \mathcal{F}_0 \subset \mathcal{F})(\exists^{\text{fin}} \mathcal{H}_0 \subset \mathcal{H})(\forall x \in \cap \mathcal{F}_0) \\
 & \quad (\exists H \in \mathcal{H}_0)(\exists y \in G(H))(\forall z \in H) \varphi.
 \end{aligned}$$

We are done on observing that $\bigcap \mathcal{F}_0 \in \mathcal{F}$ for nonempty finite subset \mathcal{F}_0 of \mathcal{F} . ▷

5.3.4. Proposition 5.3.3 enables us to describe the $\forall\exists\forall$ -cones and similar aggregates explicitly. The so-obtained descriptions are obviously bulky.

We now address the practically important constructions whose prefixes are of the types $\forall\exists$, $\forall\forall$, $\exists\forall$, and $\exists\exists$. We start with some linguistic tools for handling infinitesimals in this situation.

5.3.5. Let Ξ be a *direction*, i.e. a nonempty directed set. By idealization, there are internal elements serving as upper bounds for ${}^\circ\Xi$ in Ξ . Recall (see 4.1.6 (3)) that these bounds are *remote*, or *infinitely large* members of Ξ . Consider the standard filterbase $\mathcal{B} := \{\sigma(\xi) : \xi \in \Xi\}$ for the so-called tail filter of Ξ , where σ is the order on Ξ . The monad of the tail filter of Ξ obviously comprises the remote elements of Ξ . The following notations are current: ${}^a\Xi = \mu(\mathcal{B})$ and $\xi \approx +\infty \leftrightarrow \xi \in {}^a\Xi$.

5.3.6. Let Ξ and H be two directed sets, and let $\xi := \xi(\cdot) : H \leftrightarrow \Xi$ be a mapping. Then the following are equivalent:

- (1) $\xi({}^aH) \subset {}^a\Xi$;
- (2) $(\forall \xi \in \Xi)(\exists \eta \in H)(\forall \eta' \geq \eta)(\xi(\eta') \geq \xi)$.

◁ Indeed, (1) implies that the tail filter of Ξ is coarser than the image of the tail filter of H . This implies that each tail of Ξ includes the image of some tail of H , which is the claim of (2). ▷

5.3.7. Whenever equivalent conditions 5.3.6 (1) and 5.3.6 (2) are fulfilled, H is a *subdirection* of Ξ relative to $\xi(\cdot)$.

5.3.8. Let X be a set, and let $x := x(\cdot) : \Xi \rightarrow X$ be a net in X (in brief notations, $(x_\xi)_{\xi \in \Xi}$ or (x_ξ)). Assume further that $(y_\eta)_{\eta \in H}$ is another net in X . Say that (y_η) is a *Moore subnet* of (x_ξ) , or a *strict subnet* of (x_ξ) provided that H is a subdirection of Ξ relative to some $\xi(\cdot)$ satisfying $y_\eta = x_{\xi(\eta)}$ for all $\eta \in H$, i.e., $y = x \circ \xi$. It is worth observing that $y({}^aH) \subset x({}^a\Xi)$ by 4.1.6 (5).

5.3.9. The above-mentioned property of Moore subnets is a cornerstone of a more liberal definition of subnet whose attractive feature is a close connection with filters. Namely, a net $(y_\eta)_{\eta \in H}$ in X is a *subnet* (or a *subnet in a broader sense*) of $(x_\xi)_{\xi \in \Xi}$, provided that

$$(\forall \xi \in \Xi)(\exists \eta \in H)(\forall \eta' \geq \eta)(\exists \xi' \geq \xi)(x(\xi') = y(\eta')),$$

i.e. in the case when every tail of x contains some tail of y . In terms of monads, this reads $y({}^aH) \subset x({}^a\Xi)$ or, which is more lucid,

$$(\forall \eta \approx +\infty)(\exists \xi \approx +\infty)(y_\eta = x_\xi).$$

Speaking expressively, we say that $(x_\eta)_{\eta \in H}$ is a subset of $(x_\xi)_{\xi \in \Xi}$ (this ambiguity may lead to confusion).

It is worth observing that an arbitrary subnet is not necessarily a Moore subnet. Note also that two nets in a single set are *equivalent* provided that each one of them is a subnet of the other; i.e., if their monads coincide.

5.3.10. If \mathcal{F} is a filter on X and (x_ξ) is a net in X then we say that (x_ξ) is *subordinate* to \mathcal{F} whenever $\xi \approx \infty \leftrightarrow x_\xi \in \mu(\mathcal{F})$. In other words, (x_ξ) is subordinate to \mathcal{F} provided that the tail filter of (x_ξ) is finer than \mathcal{F} . In this event we write $x_\xi \downarrow \mathcal{F}$ by analogy with the topological notation of convergence which slightly abuses the language. Note also that if \mathcal{F} is an ultrafilter then \mathcal{F} coincides with the tail filter of every net (x_ξ) subordinate to \mathcal{F} , i.e., such a net (x_ξ) is an *ultranet*.

5.3.11. Theorem. Let $\varphi = \varphi(x, y, z)$ be a formula of Zermelo–Fraenkel set theory with no bound variables but x, y, z and z a standard set. Let \mathcal{F} be a filter on X , and let \mathcal{G} be a filter on Y . Then the following are equivalent:

- (1) $(\forall G \in \mathcal{G})(\exists F \in \mathcal{F})(\forall x \in F)(\exists y \in G) \varphi(x, y, z)$;
- (2) $(\forall x \in \mu(\mathcal{F}))(\exists y \in \mu(\mathcal{G})) \varphi(x, y, z)$;
- (3) To each net $(x_\xi)_{\xi \in \Xi}$ in X subordinate to \mathcal{F} , there are a net $(y_\eta)_{\eta \in \mathbb{H}}$ in Y subordinate to \mathcal{G} and a strict subnet $(x_{\xi(\eta)})_{\eta \in \mathbb{H}}$ of $(x_\xi)_{\xi \in \Xi}$ such that $\varphi(x_{\xi(\eta)}, y_\eta, z)$ for all $\eta \in \mathbb{H}$; in symbols,

$$(\forall x_\xi \downarrow \mathcal{F})(\exists y_\eta \downarrow \mathcal{G}) \varphi(x_{\xi(\eta)}, y_\eta, z);$$

- (4) To each set $(x_\xi)_{\xi \in \Xi}$ in X subordinate to \mathcal{F} , there are a net $(y_\eta)_{\eta \in \mathbb{H}}$ in Y subordinate to \mathcal{G} and a subnet $(x_\eta)_{\eta \in \mathbb{H}}$ of $(x_\xi)_{\xi \in \Xi}$ such that $\varphi(x_\eta, y_\eta, z)$ for all $\eta \in \mathbb{H}$; in symbols,

$$(\forall x_\xi \downarrow \mathcal{F})(\exists y_\eta \downarrow \mathcal{G}) \varphi(x_\eta, y_\eta, z);$$

- (5) To each ultranet $(x_\xi)_{\xi \in \Xi}$ in X subordinate to \mathcal{F} , there are an ultranet $(y_\eta)_{\eta \in \mathbb{H}}$ subordinate to \mathcal{G} and ultranet $(x_\eta)_{\eta \in \mathbb{H}}$ equivalent to $(x_\xi)_{\xi \in \Xi}$ such that $\varphi(x_\eta, y_\eta, z)$ for all $\eta \in \mathbb{H}$.

\triangleleft (1) \rightarrow (2): Take $x \in \mu(\mathcal{F})$. By transfer, to each standard G there is a standard F satisfying $(\forall x \in F)(\exists y \in G) \varphi(x, y, z)$. Therefore, $(\forall G \in {}^\circ\mathcal{G})(\exists y \in G) \varphi(x, y, z)$. By idealization, $(\exists y)(\forall G \in {}^\circ\mathcal{G})(y \in G) \varphi(x, y, z)$. Hence, $y \in \mu(\mathcal{G})$ and $\varphi(x, y, z)$.

(2) \rightarrow (3): Let $(x_\xi)_{\xi \in \Xi}$ be a standard net in X subordinate to \mathcal{F} . Given a standard G in \mathcal{G} and $\xi \in {}^\circ\Xi$, put

$$A_{(G, \xi)} := \{\xi' \geq \xi : (\forall \xi'' \geq \xi')(\exists y \in G) \varphi(x_{\xi''}, y, z)\}.$$

By 4.1.8, ${}^a\Xi \subset A_{(G, \xi)}$. Since $A_{(G, \xi)}$ is an internal set, from the Cauchy principle it follows that ${}^\circ A_{(G, \xi)} \neq \emptyset$. Therefore, there are standard mappings $\xi : \mathbb{H} \rightarrow \Xi$ and $y : \mathbb{H} \rightarrow Y$ on the direction $\mathbb{H} := \mathcal{G} \times \Xi$ (under the natural order) such that $\xi(\eta) \in A_{(G, \xi)}$ and $y_\eta \in G$ for $G \in \mathcal{G}$ and $\xi \in \Xi$ with $\eta = (G, \xi)$. Obviously, $\xi(\eta) \approx +\infty$ and $y_\eta \in \mu(\mathcal{G})$ for $\eta \approx +\infty$.

(3) \rightarrow (4): Obvious.

(4) \rightarrow (1): If (1) fails then, by hypothesis,

$$(\exists G \in \mathcal{G})(\forall F \in \mathcal{F})(\exists x \in F)(\forall y \in G)\neg\varphi(x, y, z).$$

Given $F \in \mathcal{F}$, choose $x_F \in F$ so that $\neg\varphi(x, y, z)$ for all $y \in G$. The so-obtained net $(x_F)_{F \in \mathcal{F}}$ in X , as well as the set G , can be considered standard by transfer. Clearly, $x_F \downarrow \mathcal{F}$. Hence, by (3) there are a direction H and a subnet $(x_\eta)_{\eta \in H}$ of $(x_F)_{F \in \mathcal{F}}$ such that we may find a subnet $(y_\eta)_{\eta \in H}$ satisfying $\varphi(x_\eta, y_\eta, z)$ for all $\eta \in H$.

By 5.3.9, if η is remote then x_η coincides with x_F for some remote F ; i.e., $x_\eta \in \mu(\mathcal{F})$. By hypothesis, $y_\eta \in \mu(\mathcal{G})$ and, moreover, $y_\eta \in G$. In this event $\varphi(x_\eta, y_\eta, z)$ and $\neg\varphi(x_\eta, y_\eta, z)$, which is a contradiction.

(1) \leftrightarrow (5): To demonstrate, note that the claim is evident in the case when \mathcal{F} and \mathcal{G} are ultrafilters. We are done on recalling that each monad is a union of ultramonads. \triangleright

5.3.12. Applications often involve specifications of 5.3.11 in which the monad of one of the filters is a singleton; i.e. the filter is *discrete*. In this event,

$$\begin{aligned} (\exists x \in \mu(\mathcal{F})) \varphi(x, y) &\leftrightarrow (\exists x_\xi \downarrow \mathcal{F}) \varphi(x_\xi, y); \\ (\forall x \in \mu(\mathcal{F})) \varphi(x, y) &\leftrightarrow (\forall x_\xi \downarrow \mathcal{F})(\exists x_\eta \downarrow \mathcal{F}) \varphi(x_\eta, y). \end{aligned}$$

5.3.13. Let $F \subset X \times Y$ be an internal correspondence from a standard set X to a standard set Y . Assume given a standard filter \mathcal{N} on X and a topology τ on Y . Put

$$\begin{aligned} \forall\forall(F) &:= * \{y' : (\forall x \in \mu(\mathcal{N}) \cap \text{dom}(F))(\forall y \approx y')(x, y) \in F\}, \\ \exists\forall(F) &:= * \{y' : (\exists x \in \mu(\mathcal{N}) \cap \text{dom}(F))(\forall y \approx y')(x, y) \in F\}, \\ \forall\exists(F) &:= * \{y' : (\forall x \in \mu(\mathcal{N}) \cap \text{dom}(F))(\exists y \approx y')(x, y) \in F\}, \\ \exists\exists(F) &:= * \{y' : (\exists x \in \mu(\mathcal{N}) \cap \text{dom}(F))(\exists y \approx y')(x, y) \in F\}, \end{aligned}$$

with $*$ symbolizing standardization and $y \approx y'$ standing for $y \in \mu(\tau(y'))$. We call $Q_1 Q_2(F)$ the $Q_1 Q_2$ -*limit* of F (here Q_k ($k := 1, 2$) is one of the quantifiers \forall or \exists).

5.3.14. It suffices in applications to restrict consideration to the case in which F is a standard correspondence on some element of \mathcal{N} and to study the $\exists\exists$ -limit and the $\forall\exists$ -limit. The former is the *limit superior* or *upper limit*; the latter is the *limit inferior* or *lower limit* of F along \mathcal{N} .

If $(x_\xi)_{\xi \in \Xi}$ is a net in the domain of F then, implying the tail filter of $(x_\xi)_{\xi \in \Xi}$, we put

$$\begin{aligned} \text{Li}_{\xi \in \Xi} F &:= \liminf_{\xi \in \Xi} := \forall \exists (F), \\ \text{Ls}_{\xi \in \Xi} F &:= \limsup_{\xi \in \Xi} F(x_\xi) := \exists \exists (F) \end{aligned}$$

and speak about *Kuratowski limits*.

5.3.15. *If F is a standard correspondence then*

$$\begin{aligned} \exists \exists (F) &= \bigcap_{U \in \mathcal{N}} \text{cl} \left(\bigcup_{x \in U} F(x) \right); \\ \forall \exists (F) &= \bigcap_{U \in \mathcal{N}^\ddot{}} \text{cl} \left(\bigcup_{x \in U} F(x) \right), \end{aligned}$$

where $\mathcal{N}^\ddot{}$ is the grill of a filter \mathcal{N} on X , i.e., the family comprising all subsets of X meeting the monad $\mu(\mathcal{N})$.

In other words,

$$\begin{aligned} \mathcal{N}^\ddot{} &= * \{U' \subset X : U' \cap \mu(\mathcal{N}) \neq \emptyset\} \\ &= \{U' \subset X : (\forall U \in \mathcal{N})(U \cap U' \neq \emptyset)\}. \end{aligned}$$

We also note the relations:

$$\begin{aligned} \forall \exists (F) &= \bigcap_{U \in \mathcal{N}^\ddot{}} \text{int} \left(\bigcup_{x \in U} F(x) \right), \\ \forall \forall (F) &= \bigcup_{U \in \mathcal{N}} \text{int} \left(\bigcap_{x \in U} F(x) \right). \end{aligned}$$

5.3.16. Theorem 5.3.11 immediately describes the limits in the language of nets.

5.3.17. *An element y lies in the $\forall \exists$ -limit of F if and only if to each net $(x_\xi)_{\xi \in \Xi}$ in $\text{dom}(F)$ subordinate to \mathcal{N} there are a subnet $(x_\eta)_{\eta \in \mathbb{H}}$ of the net $(x_\xi)_{\xi \in \Xi}$ and a net $(y_\eta)_{\eta \in \mathbb{H}}$ convergent to y such that $(x_\eta, y_\eta) \in F$ for all $\eta \in \mathbb{H}$.*

5.3.18. *An element y lies in the $\exists \exists$ -limit of F if and only if there are a net $(x_\xi)_{\xi \in \Xi}$ in $\text{dom}(F)$ subordinate to \mathcal{N} and a net $(y_\xi)_{\xi \in \Xi}$ convergent to y , such that $(x_\xi, y_\xi) \in F$ for any $\xi \in \Xi$.*

5.3.19. *If F is an internal correspondence then*

$$\forall\forall(F) \subset \exists\forall(F) \subset \forall\exists(F) \subset \exists\exists(F).$$

Moreover, $\exists\exists(F)$ and $\forall\exists(F)$ are closed sets whereas $\forall\forall(F)$ and $\exists\forall(F)$ are open sets.

◁ The claim about inclusions is obvious. By duality, it suffices to establish for definiteness only the fact that the $\forall\exists$ -limit is closed.

If V is a standard open neighborhood of a point y' of $\text{cl}(\forall\exists(F))$, then there is some y in $\forall\exists(F)$ belonging to V . Given $x \in \mu(\mathcal{N})$, find some element y'' satisfying $y'' \in \mu(\tau(y))$ and $(x, y'') \in F$. Obviously, $y'' \in V$ since V is a neighborhood of y . Therefore,

$$(\forall x \in \mu(\mathcal{N}))(\forall V \in \circ\tau(y'))(\exists y'' \in V)(x, y'') \in F.$$

By idealization, conclude that $y' \in \forall\exists(F)$. ▷

5.3.20. The above propositions make it possible to characterize the elements of many approximating or regularizing cones in the common terms of nets (see [263, 279]). Observe in particular that the Clarke cone $\text{Cl}(F, x')$ of a set F in X appears as the Kuratowski limit:

$$\text{Cl}(F, x') = \text{Li}_{\tau(x') \times \tau_{\mathbb{R}^+}(0)} \Gamma_F,$$

where Γ_F is the *homothety* associated with F , i.e.

$$(x, \alpha, h) \in \Gamma_F \leftrightarrow h \in \frac{F - x}{\alpha} \quad (x, h \in X, \alpha > 0).$$

5.3.21. Convex analysis often operates with some special variations of Kuratowski limits that involve the epigraphs of functions acting to the *extended reals* $\overline{\mathbb{R}}$. We start with recall important properties of the classical upper and lower limits.

5.3.22. *Let $f : X \rightarrow \overline{\mathbb{R}}$ be a standard function on X , and let \mathcal{F} be a standard filter on X . If $t \in \mathbb{R}$ then*

$$\begin{aligned} \sup_{F \in \mathcal{F}} \inf f(F) \leq t &\leftrightarrow (\exists x \in \mu(\mathcal{F})) \circ f(x) \leq t, \\ \inf_{F \in \mathcal{F}} \sup f(F) \leq t &\leftrightarrow (\forall x \in \mu(\mathcal{F})) \circ f(x) \leq t. \end{aligned}$$

◁ To check the first equivalence, infer by transfer and idealization that

$$\begin{aligned} \sup_{F \in \mathcal{F}} \inf f(F) \leq t &\rightarrow (\forall F \in \mathcal{F}) \inf f(F) \leq t \\ &\rightarrow (\forall F \in \mathcal{F})(\forall \varepsilon > 0) \inf f(F) < t + \varepsilon \rightarrow (\forall \varepsilon)(\forall F)(\exists x \in F) \end{aligned}$$

$$\begin{aligned}
 f(x) < t + \varepsilon &\rightarrow (\forall^{\text{st}} \varepsilon)(\forall^{\text{st}} F) (\exists x)(x \in F \wedge f(x) < t + \varepsilon) \\
 &\quad (\exists x)(\forall^{\text{st}} \varepsilon)(\forall^{\text{st}} F)(x \in F \wedge f(x) < t + \varepsilon) \\
 &\rightarrow (\exists x \in \mu(\mathcal{F}))(\forall^{\text{st}} \varepsilon > 0) f(x) < t + \varepsilon \rightarrow (\exists x \in \mu(\mathcal{F}))^\circ f(x) \leq t
 \end{aligned}$$

(this rests on 2.2.18 (3)). Observe now that $x \in \mu(\mathcal{F}) \subset F$ for a standard element F of \mathcal{F} . Hence, $\inf f(F) \leq t$ since $\inf f(F) \leq f(x) < t + \varepsilon$ for all $\varepsilon > 0$. By transfer, $\inf f(F) \leq t$ for all internal F of \mathcal{F} , which was required.

Since $-f$ and t are standard, from the above we deduce

$$\begin{aligned}
 \inf_{F \in \mathcal{F}} \sup f(F) \geq t &\leftrightarrow - \inf_{F \in \mathcal{F}} \sup f(F) \leq -t \leftrightarrow \sup_{F \in \mathcal{F}} \inf(-f)(F) \leq t \\
 &\leftrightarrow (\exists x \in \mu(\mathcal{F}))^\circ(-f(x)) \leq -t \leftrightarrow (\exists x \in \mu(\mathcal{F}))^\circ f(x) \geq t.
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 \inf_{F \in \mathcal{F}} \sup f(F) < t &\leftrightarrow \neg \left(\inf_{F \in \mathcal{F}} \sup f(F) \geq t \right) \\
 &\leftrightarrow \neg((\exists x \in \mu(\mathcal{F}))^\circ f(x) \geq t) \leftrightarrow (\forall x \in \mu(\mathcal{F}))^\circ f(x) \leq t.
 \end{aligned}$$

Finally,

$$\begin{aligned}
 \inf_{F \in \mathcal{F}} \sup f(F) \leq t &\leftrightarrow (\forall \varepsilon > 0) \inf_{F \in \mathcal{F}} \sup f(F) < t + \varepsilon \\
 &\leftrightarrow (\forall^{\text{st}} \varepsilon > 0)(\forall x \in \mu(\mathcal{F}))^\circ f(x) < t + \varepsilon \\
 &\leftrightarrow (\forall x \in \mu(\mathcal{F}))(\forall^{\text{st}} \varepsilon > 0)^\circ f(x) < t + \varepsilon \leftrightarrow (\forall x \in \mu(\mathcal{F}))^\circ f(x) \leq t,
 \end{aligned}$$

since $^\circ f(x)$ is standard. \triangleright

5.3.23. *Let X and Y be standard sets and let $f : X \times Y \rightarrow \overline{\mathbb{R}}$ be a standard function. Assume further that \mathcal{F} and \mathcal{G} are standard filters on X and Y , respectively.*

If t is a standard real t then

$$\begin{aligned}
 \sup_{G \in \mathcal{G}} \inf_{F \in \mathcal{F}} \sup_{x \in F} \inf_{y \in G} f(x, y) &\leq t \\
 &\leftrightarrow (\forall x \in \mu(\mathcal{F}))(\exists y \in \mu(\mathcal{G}))^\circ f(x, y) \leq t.
 \end{aligned}$$

\triangleleft Put $f_G(x) := \inf\{f(x, y) : y \in G\}$. Observe that f_G is a standard function whenever G is a standard set. Using successively the transfer principle, 5.3.22 and (strong) idealization, we infer that

$$\begin{aligned}
& \sup_{G \in \mathcal{G}} \inf_{F \in \mathcal{F}} \sup_{x \in F} \inf_{y \in G} f(x, y) \leq t \leftrightarrow (\forall G \in \mathcal{G}) \inf_{F \in \mathcal{F}} \sup_{x \in F} f_G(x) \leq t \\
\leftrightarrow & (\forall^{\text{st}} G \in \mathcal{G}) \inf_{F \in \mathcal{F}} \sup_{x \in F} f_G(x) \leq t \leftrightarrow (\forall^{\text{st}} G \in \mathcal{G}) (\forall x \in \mu(\mathcal{F})) \circ f_G(x) \leq t \\
& \leftrightarrow (\forall x \in \mu(\mathcal{F})) (\forall^{\text{st}} G \in \mathcal{G}) (\forall^{\text{st}} \varepsilon > 0) \inf_{y \in G} f(x, y) < t + \varepsilon \\
\rightarrow & (\forall x \in \mu(\mathcal{F})) (\forall^{\text{st}} \varepsilon > 0) (\forall^{\text{st}} G \in \mathcal{G}) (\exists y \in G) (f(x, y) < t + \varepsilon) \\
& \rightarrow (\forall x \in \mu(\mathcal{F})) (\exists y \in \mu(\mathcal{G})) (\forall^{\text{st}} \varepsilon > 0) (f(x, y) < t + \varepsilon) \\
& \rightarrow (\forall x \in \mu(\mathcal{F})) (\exists y \in \mu(\mathcal{G})) \circ f(x, y) \leq t.
\end{aligned}$$

Given an internal element $F \subset \mu(\mathcal{F})$ of \mathcal{F} and a standard element G of \mathcal{G} , the last relation yields

$$\begin{aligned}
& \sup_{x \in F} \inf_{y \in G} f(x, y) \leq t \rightarrow \inf_{F \in \mathcal{F}} \sup_{x \in F} \inf_{y \in G} f(x, y) \leq t \\
& \rightarrow (\forall^{\text{st}} G \in \mathcal{G}) \inf_{F \in \mathcal{F}} \sup_{x \in F} \inf_{y \in G} f(x, y) \leq t \\
& \rightarrow (\forall G \in \mathcal{G}) \inf_{F \in \mathcal{F}} \sup_{x \in F} \inf_{y \in G} f(x, y) \leq t,
\end{aligned}$$

which completes the proof by transfer. \triangleright

5.3.24. In view of 5.3.23 it is in common parlance to call

$$\limsup_{\mathcal{F}} \inf_{\mathcal{G}} f := \sup_{G \in \mathcal{G}} \inf_{F \in \mathcal{F}} \sup_{x \in F} \inf_{y \in G} f(x, y)$$

the *Rockafellar limit* of f .

If $f := (f_\xi)_{\xi \in \Xi}$ is a family of functions from a topological space (X, σ) to $\overline{\mathbb{R}}$, and if \mathcal{N} is a filter on Ξ , then we define the *limit inferior* or *lower limit* of f at $x' \in X$ and the *limit superior*, or *upper limit*, or the *Rockafellar limit* of f at $x \in X$ as

$$\begin{aligned}
\text{li}_{\mathcal{N}} f(x') &:= \sup_{V \in \sigma(x')} \sup_{U \in \mathcal{N}} \inf_{\xi \in U} \inf_{x \in V} f_\xi(x), \\
\text{ls}_{\mathcal{N}}(x') &:= \sup_{V \in \sigma(x')} \inf_{U \in \mathcal{N}} \sup_{\xi \in U} \inf_{x \in V} f_\xi(x).
\end{aligned}$$

These last limits are often referred to as *epilimits*. The idea behind this is clear from the following easy proposition.

5.3.25. *The upper and lower limits of a family of epigraphs are the epigraphs of the respective limits of the family of functions under consideration.*

5.4. Approximation Given a Set of Infinitesimals

In this section we analyze the classical approximating cones of Clarke type by elaborating the contribution of infinitesimals to their definition. This analysis enables us to single out new analogs of tangent cones as well as new descriptions for the Clarke cone.

5.4.1. We again consider a real vector space X furnished with a linear topology σ and a nearvector topology τ . Assume that F is a set in X and x' is a point in F . In line with 5.2, these objects are considered standard.

We start with fixing a positive infinitesimal real number α and putting

$$\begin{aligned} \text{Ha}_\alpha(F, x') &:= {}^* \{h' \in X : (\forall x \approx_\tau x', x \in F)(\forall h \approx_\tau h')(x + \alpha h \in F)\}, \\ \text{In}_\alpha(F, x') &:= {}^* \{h' \in X : (\exists h \approx_\tau h')(\forall x \approx_\sigma x', x \in F)(x + \alpha h \in F)\}, \\ \text{Cl}_\alpha(F, x') &:= {}^* \{h' \in X : (\forall x \approx_\sigma x', x \in F)(\exists h \approx_\tau h')(x + \alpha h \in F)\}, \end{aligned}$$

where $*$ stands as usual for standardization.

We now consider a nonempty and, generally speaking, external set of positive infinitesimals Λ and put

$$\begin{aligned} \text{Ha}_\Lambda(F, x') &:= {}^* \bigcap_{\alpha \in \Lambda} \text{Ha}_\alpha(F, x'), \\ \text{In}_\Lambda(F, x') &:= {}^* \bigcap_{\alpha \in \Lambda} \text{In}_\alpha(F, x'), \\ \text{Cl}_\Lambda(F, x') &:= {}^* \bigcap_{\alpha \in \Lambda} \text{Cl}_\alpha(F, x'). \end{aligned}$$

We will pursue the same policy as regards notation for other types of the approximations we introduce. By way of example, it is worth emphasizing that by definition

$$\begin{aligned} h' \in \text{In}_\Lambda(F, x') \\ \leftrightarrow (\forall \alpha \in \Lambda)(\exists h \approx_\tau h')(\forall x \approx_\sigma x', x \in F)(x + \alpha h \in F) \end{aligned}$$

for every standard point h' of X .

If Λ is the monad of the corresponding standard filter \mathcal{F}_Λ , where $\mathcal{F}_\Lambda := {}^* \{A \subset \mathbb{R} : A \supset \Lambda\}$ then, say, for $\text{Cl}_\Lambda(F, x')$ we have

$$\text{Cl}_\Lambda(F, x') = \bigcap_{V \in \mathcal{N}} \bigcup_{\substack{U \in \sigma(x') \\ A \in \mathcal{F}_\Lambda}} \bigcap_{\substack{x \in F \cap U \\ \alpha \in A, \alpha > 0}} \left(\frac{F - x}{\alpha} + V \right).$$

If Λ is not a monad (for instance, a singleton) then the implicit form of $\text{Cl}_\Lambda(F, x')$ is determined by the particular choice of the model of analysis we deal with. It is worth emphasizing that the monad of the ultrafilter $\mathcal{U}(\alpha) := \ast\{A \subset \mathbb{R} : \alpha \in A\}$ reduces in no way to the initial infinitesimal α ; i.e., the set $\text{Cl}_\alpha(F, x')$ is, generally speaking, larger than $\text{Cl}_{\mu(\mathcal{U}(\alpha))}(F, x')$. At the same time, the above approximations happily enjoy many advantageous properties inherent to Clarke cones. Elaborating the last claim, we will suppose as before in 5.2 that the mapping $(x, \beta, h) \mapsto x + \beta h$ from $(X \times \mathbb{R} \times X, \sigma \times \tau_{\mathbb{R}} \times \tau)$ to (X, σ) is continuous at zero, which amounts to the inclusion $\mu(\sigma) + \mu(\mathbb{R}_+) \cdot \mu(\tau) \subset \mu(\sigma)$ on assuming the standard environment.

5.4.2. Theorem. *If Λ is a set of positive infinitesimals then*

(1) $\text{Ha}_\Lambda(F, x')$, $\text{In}_\Lambda(F, x')$, and $\text{Cl}_\Lambda(F, x')$ are semigroups and

$$\text{Ha}(F, x') \subset \text{Ha}_\Lambda(F, x') \subset \text{In}_\Lambda(F, x') \subset \text{Cl}_\Lambda(F, x') \subset \text{K}(F, x'),$$

$$\text{Cl}(F, x') \subset \text{Cl}_\Lambda(F, x');$$

(2) *If Λ is internal then $\text{Ha}_\Lambda(F, x')$ is τ -open;*

(3) $\text{Cl}_\Lambda(F, x')$ is τ -closed; moreover, if F is convex then we have

$$\text{K}(F, x') = \text{Cl}_\Lambda(F, x')$$

whenever $\sigma = \tau$;

(4) *If $\sigma = \tau$ then*

$$\text{Cl}_\Lambda(F, x') = \text{Cl}_\Lambda(\text{cl}(F), x');$$

(5) *The Rockafellar formula holds*

$$\text{Ha}_\Lambda(F, x') + \text{Cl}_\Lambda(F, x') = \text{Ha}_\Lambda(F, x');$$

(6) *If x' is a τ -boundary point of F then*

$$\text{Ha}_\Lambda(F, x') = -\text{Ha}_\Lambda(F', x'),$$

with $F' := (X - F) \cup \{x'\}$.

◁ (1): We must only check that $\text{In}_\alpha(F, x')$ is a semigroup. If h' and h'' are standard and belong to $\text{In}_\lambda(F, x')$ then there is some $h_1 \approx_\sigma h'$ such that $x'' := x + \alpha h_1 \in F$ for all $x \in F$, $x \approx_\sigma x'$. By hypothesis, there is also some $h_2 \approx_\tau h''$ such that $x'' + \alpha h_2 \in F$ whenever $x'' \approx_\sigma x$. Finally, $h_1 + h_2 \approx_\tau h' + h''$ and, moreover, this $h_1 + h_2$ “serves” to the membership $h' + h'' \in \text{In}_\lambda(F, x')$.

If $h' \in \text{Cl}_\alpha(F, x')$ and h' is standard then $x' + \alpha h \in F$ and $h \approx_\tau h'$, which implies $h' \in \text{K}(F, x')$. The rest of the inclusions in (1) are obvious.

(2): If h' is a standard element of $\text{Ha}_\Lambda(F, x')$ then

$$(\forall x \approx_\sigma x', x \in F)(\forall h \approx_\tau h')(\forall \alpha \in \Lambda)(x + \alpha h \in F).$$

Taking 5.3.2 into account and using the fact that Λ is an internal set, we deduce:

$$(\exists^{\text{st}} V \in \mathcal{N}_\tau)(\exists^{\text{st}} U \in \sigma(x'))(\forall x \in U \cap F)(\forall h \in h' + V)(\forall \alpha \in \Lambda)(x + \alpha h \in F).$$

Choose standard neighborhoods $V_1, V_2 \in \mathcal{N}_\tau$ so that $V_1 + V_2 \subset V$. Then

$$(\forall x \in U \cap F)(\forall h \in h'' + V_2)(\forall \alpha \in \Lambda)(x + \alpha h \in F)$$

for all standard $h'' \in h' + V_1$; i.e. $h'' \in \text{Ha}_\Lambda(F, x')$ for all $h'' \in h' + V_1$.

(3): Assume now that h' is a standard element of $\text{cl}_\tau(\text{Cl}_\Lambda(F, x'))$. Take an arbitrary standard neighborhood V of h' and again choose some standard $V_1, V_2 \in \mathcal{N}_\tau$ that satisfy the condition $V_1 + V_2 \subset V$. By definition, there is some h'' in $\text{Cl}_\Lambda(F, x')$ such that $h'' \in h' + V_1$. By 5.4.1 and 5.3.2,

$$(\forall \alpha \in \Lambda)(\exists^{\text{st}} U \in \sigma(x'))(\forall x \in F \cap U)(\exists h \in h'' + V_2)(x + \alpha h \in F).$$

Moreover, $h \in h'' + V_2 \subset h' + V_1 + V_2 \subset h' + V$. In other words,

$$(\forall^{\text{st}} V \in \mathcal{N}_\tau)(\forall \alpha \in \Lambda)(\exists^{\text{st}} U \in \sigma(x'))(\forall x \in F \cap U)(\exists h \in h' + V)(x + \alpha h \in F).$$

Therefore, $h' \in \text{Cl}_\alpha(F, x')$ for all $\alpha \in \Lambda$, i.e., $h' \in \text{Cl}_\Lambda(F, x')$.

Assume now that $h' \in \text{Fd}_\Lambda(F, x')$ and h' is standard. Then $x' + \alpha' h' \in F$ for some standard $\alpha' > 0$ by transfer. If $x \approx_\sigma x'$ and $x \in F$ then $(x - x')/\alpha' \approx_\sigma 0$. Putting $h := h' + (x - x')/\alpha'$ obtain $h \approx_\tau h'$ and, moreover, $x + \alpha' h \in F$. Since F is convex, conclude that $x + (0, \alpha']h \subset F$. In particular, $x + \Lambda h \subset F$. Hence,

$$(\forall x \approx_\sigma x', x \in F)(\forall \alpha \in \Lambda)(\exists h \approx_\tau h')(x + \alpha h \in F);$$

i.e., $h' \in \text{Cl}_\Lambda(F, x')$. Hence,

$$\text{Fd}(F, x') \subset \text{Cl}_\Lambda(F, x') \subset \text{K}(F, x') \subset \text{cl}(\text{Fd}(F, x')).$$

Since $\text{Cl}_\Lambda(F, x')$ is τ -closed; therefore, $\text{K}(F, x') = \text{Cl}_\Lambda(F, x')$.

(4): The proof proceeds along the same lines as in 5.2.11.

(5): Given standard $k' \in \text{Ha}_\Lambda(F, x')$ and $h' \in \text{Cl}_\Lambda(F, x')$, for all $\alpha \in \Lambda$ and all $x \in F$ satisfying $x \approx_\sigma x'$, choose h from the conditions $h \approx_\tau h'$ and $x + \alpha h \in F$. We then obtain

$$\begin{aligned}
x + \alpha(h' + k' + \mu(\tau)) &= x + \alpha h + \alpha(k' + (h - h') + \mu(\tau)) \\
&\subset (x + \mu(\sigma)) \cap F + \alpha(k' + \mu(\tau) + \mu(\tau)) \\
&\subset (x + \mu(\sigma)) \cap F + \alpha(k' + \mu(\tau)) \subset F,
\end{aligned}$$

which means that $h' + k'$ belongs to $\text{Ha}_\Lambda(F, x')$.

(6): Let $-h \notin \text{Ha}_\Lambda(F', x')$. Then to some α in Λ there is an element $h \approx_\tau h'$ satisfying $x - \alpha h \in F$ for an appropriate $x \approx_\sigma x'$, $x \in F$. If, nevertheless, $h \in \text{Ha}_\Lambda(F, x')$ then, in particular, $h \in \text{Ha}_\alpha(F, x')$ and $x = (x - \alpha h) + \alpha h \in F$, since $x - \alpha h \approx_\sigma x$. Hence, $x \in F \cap F'$, i.e. $x = x'$. Moreover, $(x' - \alpha h) + \alpha(h + \mu(\tau)) \subset F$, since $h + \mu(\tau) \subset \mu(\tau(h'))$. Therefore, x' is a τ -interior point of F , which contradicts the assumption. Hence, $h \notin \text{Ha}_\Lambda(F, x')$, which ensures the inclusion $-\text{Ha}_\Lambda(F, x') \subset \text{Ha}_\Lambda(F', x')$. Substituting $F = (F')'$ for F' in the above considerations, we complete the proof of (6). \triangleright

5.4.3. It is important to note that the above analogs of the Hadamard and Clarke cones are convex in many situations, as shown by the following propositions.

5.4.4. Let τ be a vector topology, and $t\Lambda \subset \Lambda$ for some standard $t \in (0, 1)$. Then $\text{Cl}_\Lambda(F, x')$ is a convex cone. If, in addition, Λ is an internal set then $\text{Ha}_\Lambda(F, x')$ is a convex cone too.

\triangleleft Consider only the case of $\text{Ha}_\Lambda(F, x')$ and take a standard h in $\text{Ha}_\Lambda(F, x')$. By 5.4.2(2) $\text{Ha}_\Lambda(F, x')$ is open in the topology τ . Moreover, $th \in \text{Ha}_\Lambda(F, x')$, where t is the standard positive real of the hypothesis. \triangleright

5.4.5. Let $t\Lambda \subset \Lambda$ for all standard $t \in (0, 1)$. Then $\text{Cl}_\Lambda(F, x')$, $\text{In}_\Lambda(F, x')$, and $\text{Ha}_\Lambda(F, x')$ are convex cones.

\triangleleft For definiteness, we settle the case of $\text{Cl}_\Lambda(F, x')$. Let h' be a standard vector of $\text{Cl}_\Lambda(F, x')$, and let $0 < t < 1$ be a standard real. Take $x \approx_\sigma x'$, $x \in F$ and $\alpha \in \Lambda$. Considering x and $t\alpha \in \Lambda$, choose an element h , such that $h \approx_\tau h'$ and $x + \alpha h \in F$. Since $th \approx_\tau th'$ by 5.1.7; therefore, $th' \in \text{Cl}_\alpha(F, x')$. In other words, $(0, 1)\text{Cl}_\Lambda(F, x') \subset \text{Cl}_\Lambda(F, x')$ by transfer. We are done on recalling 5.4.2(1). \triangleright

5.4.6. A set Λ is *representative* provided that $\text{Ha}_\Lambda(F, x')$ and $\text{Cl}_\Lambda(F, x')$ are convex cones. Propositions 5.4.4 and 5.4.5 exhibit examples of representative Λ 's.

5.4.7. Let $f : X \rightarrow \overline{\mathbb{R}}$ be a function acting to the extended reals. Given a positive infinitesimal α , a point x' in $\text{dom}(f)$, and a vector $h' \in X$, put

$$\begin{aligned}
f(\text{Ha}_\alpha)(x')(h') &:= \inf\{t \in \mathbb{R} : (h', t) \in \text{Ha}_\alpha(\text{epi}(f), (x', f(x')))\}, \\
f(\text{In}_\alpha)(x')(h') &:= \inf\{t \in \mathbb{R} : (h', t) \in \text{In}_\alpha(\text{epi}(f), (x', f(x')))\}, \\
f(\text{Cl}_\alpha)(x')(h') &:= \inf\{t \in \mathbb{R} : (h', t) \in \text{Cl}_\alpha(\text{epi}(f), (x', f(x')))\}.
\end{aligned}$$

The derivatives $f(\text{Ha}_\alpha)$, $f(\text{In}_\alpha)$, and $f(\text{Cl}_\alpha)$ are introduced in a natural manner.

The function $f(\text{Cl}) := f(\text{Cl}_{\mu(\mathbb{R}_+)})$ is called the *Rockafellar derivative* and denoted by f^\uparrow . Thus,

$$f_\alpha^\uparrow(x') := f(\text{Cl}_\alpha)(x'), \quad f_\Lambda^\uparrow(x') := f(\text{Cl}_\Lambda)(x').$$

If τ is discrete then $\text{Ha}_\Lambda(F, x') = \text{In}_\Lambda(F, x') = \text{Cl}_\Lambda(F, x')$. In this case the Rockafellar derivative is called the *Clarke derivative*. The following notation is common:

$$f_\alpha^\circ(x') := f_\alpha^\uparrow(x'), \quad f_\Lambda^\circ(x') := f_\Lambda^\uparrow(x').$$

If $\Lambda = \mu(\mathbb{R}_+)$ then we omit any indications of Λ .

Considering epiderivatives, we assume that the space $X \times \mathbb{R}$ is endowed with the conventional product topologies $\sigma \times \tau_{\mathbb{R}}$ and $\tau \times \tau_{\mathbb{R}}$, where $\tau_{\mathbb{R}}$ is the conventional topology on \mathbb{R} . It is sometimes convenient to furnish $X \times \mathbb{R}$ with the couple of the topologies $\sigma \times \tau_0$ and $\tau \times \tau_{\mathbb{R}}$, where τ_0 is the trivial topology in \mathbb{R} . Using these topologies, we speak about the *Clarke and Rockafellar derivatives along effective domain* $\text{dom}(f)$ and add the index d to the notation: f_d° , $f_{\Lambda, d}^\uparrow$, etc.

5.4.8. *The following hold:*

$$\begin{aligned} & f_\alpha^\uparrow(x')(h') \leq t' \\ \Leftrightarrow & (\forall x \approx_\sigma x', t \approx f(x'), t \geq f(x)) (\exists h \approx_\tau h')^\circ ((f(x + \alpha h) - t)/\alpha) \leq t'; \\ & f_\alpha^\circ(x')(h') < t' \\ \Leftrightarrow & (\forall x \approx_\sigma x', t \approx f(x'), t \geq f(x)) (\forall h \approx_\tau h')^\circ ((f(x + \alpha h) - t)/\alpha) < t'; \\ & f_{\alpha, d}^\uparrow(x')(h') \leq t' \\ \Leftrightarrow & (\forall x \approx_\sigma x', x \in \text{dom}(f)) (\exists h \approx_\tau h')^\circ ((f(x + \alpha h) - t)/\alpha) \leq t'; \\ & f_{\alpha, d}^\circ(x')(h') < t' \\ \Leftrightarrow & (\forall x \approx_\sigma x', x \in \text{dom}(f)) (\forall h \approx_\tau h')^\circ ((f(x + \alpha h) - t)/\alpha) < t'. \end{aligned}$$

◁ To prove, appeal to 2.2.18 (3). ▷

5.4.9. *If f is a lower semicontinuous function then*

$$\begin{aligned} & f_\alpha^\uparrow(x')(h') \leq t' \\ \Leftrightarrow & (\forall x \approx_\sigma x', f(x) \approx f(x')) (\exists h \approx_\tau h') \left(\frac{f(x + \alpha h) - f(x)}{\alpha} \right) \leq t'; \\ & f_\alpha^\circ(x')(h') < t' \end{aligned}$$

$$\leftrightarrow (\forall x \approx_{\sigma} x', f(x) \approx f(x')) (\forall h \approx_{\tau} h') \left(\frac{f(x + \alpha h) - f(x)}{\alpha} \right) < t'.$$

◁ Only the implications from left to right need checking. We settle the first case since both proofs are identical. Since f is lower semicontinuous; therefore, $x' \approx_{\sigma} x \rightarrow {}^{\circ}f(x) \geq f(x')$. Consequently, if x and t satisfy $t \approx f(x')$ and $t \geq f(x)$ then ${}^{\circ}t \geq {}^{\circ}f(x) \geq f(x') = {}^{\circ}t$. In other words, ${}^{\circ}f(x) = f(x')$ and $f(x) \approx f(x')$. Choosing h by hypothesis, come to the conclusion

$${}^{\circ}(\alpha^{-1}(f(x + \alpha h) - t)) \leq {}^{\circ}(\alpha^{-1}(f(x + \alpha h) - f(x))) \leq t',$$

which completes the proof (see 4.2.7). ▷

5.4.10. *If f is a continuous function then*

$$f_{\Lambda, d}^{\uparrow}(x') = f_{\Lambda}^{\uparrow}(x'), \quad f_{\Lambda, d}^{\circ}(x') = f_{\Lambda}^{\circ}(x').$$

◁ It suffices to note that the continuity of f at a standard point means in symbols that $(x \approx_{\sigma} x', x \in \text{dom}(f)) \rightarrow f(x) \approx f(x')$. ▷

5.4.11. Theorem. *Let Λ be a monad. Then*

(1) *If f is a lower semicontinuous function then*

$$f_{\Lambda}^{\uparrow}(x')(h') = \limsup_{\substack{x \rightarrow_f x' \\ \alpha \in \mathcal{F}_{\Lambda}}} \inf_{h \rightarrow h'} \frac{f(x + \alpha h) - f(x)}{\alpha},$$

$$f_{\Lambda}^{\circ}(x')(h') = \limsup_{\substack{x \rightarrow_f x' \\ \alpha \in \mathcal{F}_{\Lambda}}} \frac{f(x + \alpha h') - f(x)}{\alpha},$$

where $x \rightarrow_f x'$ means that $x \rightarrow_{\sigma} x'$ and $f(x) \rightarrow f(x')$;

(2) *If f is a continuous function then*

$$f_{\Lambda, d}^{\uparrow}(x')(h') = \limsup_{\substack{x \rightarrow x' \\ \alpha \in \mathcal{F}_{\Lambda}}} \inf_{h \rightarrow h'} \frac{f(x + \alpha h) - f(x)}{\alpha},$$

$$f_{\Lambda, d}^{\circ}(x')(h') = \limsup_{\substack{x \rightarrow x' \\ \alpha \in \mathcal{F}_{\Lambda}}} \frac{f(x + \alpha h') - f(x)}{\alpha}.$$

◁ To prove, recall 5.3.23 together with 5.4.9 and 5.4.10. ▷

5.4.12. Theorem. *Let Λ be a representative set of infinitesimals. Then*

(1) *If f is a directionally Lipschitz mapping at a point x' , i.e.*

$$\text{Ha}(\text{epi}(f), (x', f(x'))) \neq \emptyset;$$

then

$$f_{\Lambda}^{\uparrow}(x') = f_{\Lambda}^{\circ}(x').$$

If, moreover, f is continuous at x' then

$$f_{\Lambda}^{\uparrow}(x') = f_{\Lambda,d}^{\uparrow}(x') = f_{\Lambda,d}^{\circ}(x') = f_{\Lambda}^{\circ}(x');$$

(2) *If f is an arbitrary function and the Hadamard cone of $\text{dom}(f)$ at x' is nonempty, i.e. $\text{Ha}(\text{dom}(f), x') \neq \emptyset$; then*

$$f_{\Lambda,d}^{\uparrow}(x') = f_{\Lambda,d}^{\circ}(x').$$

◁ The proof of both statements sought proceeds along the same lines as in Theorem 5.4.2. We will elaborate only the case of a directionally Lipschitz f .

Put $\mathcal{A} := \text{epi}(f)$ and $a' := (x', f(x'))$. By hypothesis, the sets $\text{Cl}_{\Lambda}(\mathcal{A}, a')$ and $\text{Ha}_{\Lambda}(\mathcal{A}, a')$ are convex cones. Moreover, $\text{Ha}_{\Lambda}(\mathcal{A}, a') \supset \text{Ha}(\mathcal{A}, a')$ and so

$$\text{int}_{\tau \times \tau_{\mathbb{R}}} \text{Ha}_{\Lambda}(\mathcal{A}, a') \neq \emptyset.$$

By the Rockafellar formula,

$$\text{cl}_{\tau \times \tau_{\mathbb{R}}}(\text{Ha}_{\Lambda}(\mathcal{A}, a')) = \text{Cl}_{\Lambda}(\mathcal{A}, a'),$$

which completes the proof. ▷

5.4.13. Theorem. *Let $f_1, f_2 : X \rightarrow \overline{\mathbb{R}}$ be arbitrary functions. Take $x' \in \text{dom}(f_1) \cap \text{dom}(f_2)$. Then*

$$(f_1 + f_2)_{\Lambda,d}^{\uparrow}(x') \leq (f_1)_{\Lambda,d}^{\uparrow}(x') + (f_2)_{\Lambda,d}^{\circ}(x').$$

If, moreover, f_1 and f_2 are continuous at x' then

$$(f_1 + f_2)_{\Lambda}^{\uparrow}(x') \leq (f_1)_{\Lambda}^{\uparrow}(x') + (f_2)_{\Lambda}^{\circ}(x').$$

◁ Choose a standard element h' so that

$$h' \in \text{dom}((f_2)_{\Lambda,d}^{\circ}) \cap \text{dom}((f_1)_{\Lambda,d}^{\uparrow}).$$

If there is no such an h' the sought estimates are obvious.

Take $t' \geq (f_1)_{\Lambda,d}^{\uparrow}(x')(h')$ and $s' > (f_2)_{\Lambda,d}^{\circ}(x')(h')$. By 5.4.8, to all $x \approx_{\sigma} x'$, $x \in \text{dom}(f_1) \cap \text{dom}(f_2)$, and $\alpha \in \Lambda$ there is an element h such that $h \approx_{\tau} h'$ and, moreover,

$$\delta_1 := \circ((f_1(x + \alpha h) - f_1(x))/\alpha) \leq t';$$

$$\delta_2 := \circ((f_2(x + \alpha h) - f_2(x))/\alpha) < s'.$$

Hence, $\delta_1 + \delta_2 < t' + s'$, which gives (1).

If f_1 and f_2 are continuous at x , then we are done on recalling 5.4.10. ▷

5.4.14. In closing this section, we address some special presentations of the Clarke cone in finite-dimensional space that rest on the following remarkable result.

5.4.15. Cornet Theorem. *If the ambient space is finite-dimensional then the Clarke cone is the Kuratowski limit of contingencies:*

$$\text{Cl}(F, x') = \text{Li}_{\substack{x \rightarrow x' \\ x \in F}} \text{K}(F, x).$$

5.4.16. *Let Λ be an external set of strictly positive infinitesimals, containing an internal vanishing sequence. Then*

$$\text{Cl}_\Lambda(F, x') = \text{Cl}(F, x').$$

◁ By transfer, we may assume the standard environment.

Since the inclusion $\text{Cl}_\Lambda(F, x') \supset \text{Cl}(F, x')$ is obvious, take a standard point h' in $\text{Cl}_\Lambda(F, x')$ and establish that h' lies in the Clarke cone $\text{Cl}(F, x')$. By 5.3.13,

$$\text{Li}_{\substack{x \rightarrow x' \\ x \in F}} \text{K}(F, x) = \{h' : (\forall x \approx x', x \in F)(\exists h \approx h') h \in \text{K}(F, x)\}.$$

Consequently, if $x \approx x'$ and $x \in F$ then there is an element h in $\text{K}(F, x)$ infinitely close to h' . If (α_n) is a vanishing sequence in Λ then, by hypothesis,

$$(\forall n \in \mathbb{N})(\exists h_n)(x + \alpha_n h_n \in F \wedge h_n \approx h').$$

For every standard $\varepsilon > 0$ and the conventional norm $\|\cdot\|$ in \mathbb{R}^n we have $\|h_n - h'\| \leq \varepsilon$. Since bounded sets are precompact in finite dimensions, there are sequences $(\bar{\alpha}_n)$ and (\bar{h}_n) such that

$$\bar{\alpha}_n \rightarrow 0, \quad \bar{h}_n \rightarrow \bar{h}, \quad \|\bar{h} - h'\| \leq \varepsilon, \quad x + \bar{\alpha}_n \bar{h}_n \in F \quad (n \in \mathbb{N}).$$

By idealization, infer that there are some sequences $(\bar{\alpha}_n)$ and (\bar{h}_n) serving simultaneously to every standard positive real ε . Obviously, the corresponding limit vector h is infinitely close to h' , and, at the same time, $h \in \text{K}(F, x)$ by the definition of contingency. ▷

5.4.17. We may take as Λ in 5.4.16 the monad of an arbitrary vanishing filter, for instance, the tail filter of a vanishing standard sequence (α_n) of strictly positive reals. We will list the characteristics of the Clarke cone pertaining to this case and supplementing those above. In formulation we let the symbol $d_F(x)$ stand for the distance from a point x to a set F .

5.4.18. Theorem. *Given a vanishing sequence (α_n) of strictly positive reals, the following are equivalent:*

- (1) $h' \in \text{Cl}(F, x')$;
- (2) $\limsup_{\substack{x \rightarrow x' \\ n \rightarrow \infty}} \frac{d_F(x + \alpha_n h') - d_F(x)}{\alpha_n} \leq 0$;
- (3) $\limsup_{x \rightarrow x'} \limsup_{n \rightarrow \infty} \alpha_n^{-1} (d_F(x + \alpha_n h') - d_F(x)) \leq 0$;
- (4) $\limsup_{\substack{x \rightarrow x' \\ x \in F}} \limsup_{n \rightarrow \infty} \alpha_n^{-1} d_F(x + \alpha_n h') = 0$;
- (5) $\limsup_{x \rightarrow x'} \liminf_{n \rightarrow \infty} \alpha_n^{-1} (d_F(x + \alpha_n h') - d_F(x)) \leq 0$;
- (6) $\lim_{\substack{x \rightarrow x' \\ x \in F}} \liminf_{n \rightarrow \infty} \frac{d_F(x + \alpha_n h')}{\alpha_n} = 0$.

◁ Observe first of all that if $\alpha > 0$ then

$$\circ(\alpha^{-1} d_F(x + \alpha h')) = 0 \leftrightarrow (\exists h \approx h')(x + \alpha h \in F),$$

with $\circ t$ standing as usual for the standard part of a real t .

Indeed, to demonstrate the implication \leftarrow , put $y := x + \alpha h'$ and find

$$d_F(x + \alpha h')/\alpha = \|x + \alpha h' - y\|/\alpha \leq \|h - h'\|.$$

Checking the reverse implication, invoke the idealization principle and successively deduce:

$$\begin{aligned} \circ(\alpha^{-1} d_F(x + \alpha h')) = 0 &\rightarrow (\forall^{\text{st}} \varepsilon > 0) d_F(x + \alpha h')/\alpha < \varepsilon \\ &\rightarrow (\forall^{\text{st}} \varepsilon > 0) (\exists y \in F) \|x + \alpha h' - y\|/\alpha < \varepsilon \\ &\rightarrow (\exists y \in F) (\forall^{\text{st}} \varepsilon > 0) \|h' - (y - x)/\alpha\| < \varepsilon \\ &\rightarrow (\exists y \in F) \|h - (y - x)/\alpha\| \approx 0. \end{aligned}$$

Putting $h := (y - x)/\alpha$, note that $h \approx h'$, and $x + \alpha h \in F$.

We now proceed to proving the claims. Since the implications (3) \rightarrow (4) \rightarrow (6) and (3) \rightarrow (5) \rightarrow (6) are obvious, we will establish only that (1) \rightarrow (2) \rightarrow (3) and (6) \rightarrow (1).

(1) \rightarrow (2): Assuming the standard environment, take $x \approx x'$ and $N \approx +\infty$. Choose $x'' \in F$ so that $\|x - x''\| \leq d_F(x') + \alpha_N^2$. Since

$$d_F(x + \alpha_N h') - d_F(x'' + \alpha_N h') \leq \|x - x''\|;$$

therefore,

$$\begin{aligned} & (d_F(x + \alpha_N h') - d_F(x))/\alpha_N \\ & \leq (d_F(x'' + \alpha_N h') + \|x - x''\| - d_F(x))/\alpha_N \\ & \leq d_F(x'' + \alpha_N h')/\alpha_N + \alpha_N. \end{aligned}$$

Considering that $h' \in \text{Cl}(F, x')$, as well as the choice of x'' and N , we have $x'' + \alpha_N h' \in F$ for some $h \approx h'$. Therefore, ${}^\circ(d_F(x'' + \alpha_N h')/\alpha_N) = 0$. Hence,

$$(\forall x \approx x')(\forall N \approx +\infty) {}^\circ(\alpha_N^{-1}(d_F(x + \alpha_N h') - d_F(x))) \leq 0.$$

By 5.3.22, this is a nonstandard reformulation of (2).

(2) \rightarrow (3): It suffices to observe that if $f : U \times V \rightarrow \overline{\mathbb{R}}$ is an extended function, \mathcal{F} is a filter on U , and \mathcal{G} is a filter on V then

$$\begin{aligned} & \limsup_{\mathcal{F}} \limsup_{\mathcal{G}} f(x, y) \leq t \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F})) {}^\circ \limsup_{\mathcal{G}} f(x, y) \leq t \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\forall^{\text{st}} \varepsilon > 0) \inf_{G \in \mathcal{G}} \sup_{y \in G} f(x, y) < t + \varepsilon \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\forall^{\text{st}} \varepsilon > 0)(\exists G \in \mathcal{G}) \sup_{y \in G} f(x, y) < t + \varepsilon \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\exists G \in \mathcal{G})(\forall^{\text{st}} \varepsilon > 0) \sup_{y \in G} f(x, y) < t + \varepsilon \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\exists G \in \mathcal{G})(\forall^{\text{st}} \varepsilon > 0) \sup_{y \in G} f(x, y) \leq t + \varepsilon \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\exists G \in \mathcal{G})(\forall y \in G) {}^\circ f(x, y) \leq t. \end{aligned}$$

As usual, $\mu(\mathcal{F})$ stands for the monad of \mathcal{F} .

(6) \rightarrow (1): Using the previous notation, observe first that

$$\begin{aligned} & \limsup_{\mathcal{F}} \liminf_{\mathcal{G}} f(x, y) \leq t \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F})) \sup_{G \in \mathcal{G}} \inf_{y \in G} f(x, y) \leq t \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\forall^{\text{st}} \varepsilon > 0)(\forall G \in \mathcal{G}) \inf_{y \in G} f(x, y) \leq t + \varepsilon \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\forall G \in \mathcal{G})(\forall^{\text{st}} \varepsilon > 0) \inf_{y \in G} f(x, y) < t + \varepsilon \\ & \leftrightarrow (\forall x \in \mu(\mathcal{F}))(\forall G \in \mathcal{G})(\forall^{\text{st}} \varepsilon > 0)(\exists y \in G) f(x, y) < t + \varepsilon \end{aligned}$$

$$\leftrightarrow (\forall x \in \mu(\mathcal{F}))(\forall G \in \mathcal{G})(\exists y \in G) \circ f(x, y) \leq t.$$

By hypothesis, the above implies

$$(\forall x \approx x', x \in F)(\forall n)(\exists N \geq n) \circ (\alpha_N^{-1} d_F(x + \alpha_N h')) = 0.$$

In other words, $x + \alpha_N h_N \in F$ for some h_N satisfying $h_N \approx h'$.

Proceeding along the same lines as in 5.4.16, conclude that h' lies in the lower Kuratowski limit of the contingencies of F at the points close to x' , i.e. in the Clarke cone $\text{Cl}(F, x')$. \triangleright

5.5. Approximation to Composites

We now proceed to studying the tangents of Clarke type to composites of correspondences. To this end we have to start with some topological preliminaries to open and nearly open operators.

5.5.1. As before, we consider a vector space X furnished with two topologies σ_X and τ_X as well as another vector space Y with topologies σ_Y and τ_Y . Let T be a linear operator from X to Y . We will first address the interplay between the approximating sets to F at a point x' , with $F \subset X$, and to the image $T(F)$ of F under T at Tx' .

5.5.2. *The following hold:*

(1) *The inclusion*

$$T(\mu(\sigma_X(x')) \cap F) \supset \mu(\sigma_Y(Tx')) \cap T(F)$$

amounts to the relative preopenness (ρ_-) condition (with parameters T , F , and x');

$$(\forall U \in \sigma_X(x'))(\exists V \in \sigma_Y(Tx')) T(U \cap F) \supset V \cap T(F);$$

(2) *Condition (ρ_-) combined with the requirement that T is a continuous mapping from (X, σ_X) to (Y, σ_Y) amounts to the relative openness or (ρ) condition:*

$$T(\mu(\sigma_X(x')) \cap F) = \mu(\sigma_Y(Tx')) \cap T(F);$$

(3) *The operator T satisfies the relative quasiopenness or ($\bar{\rho}$) condition*

$$\begin{aligned} &(\forall U \in \sigma_X(x'))(\exists V \in \sigma_Y(Tx')) \\ &(\text{cl}_{\tau_Y}(T(U \cap F)) \supset V \cap T(F)) \end{aligned}$$

if and only if

$$(\forall W \in \mathcal{N}_{\tau_Y})(T(\mu(\sigma_X(x')) \cap F) + W) \supset \mu(\sigma_Y(Tx')) \cap T(F)).$$

◁ Claims (1) and (2) ensue from specialization of 5.3.2. To prove (3), we first put

$$\begin{aligned} \mathcal{A} &:= T(\sigma_X(x') \cap F), \quad \mathcal{B} := \sigma_Y(Tx') \cap T(F), \\ \mathcal{N} &:= \{N \subset Y^2 : (\exists W \in \mathcal{N}_{\tau_Y}) N \supset \{(y_1, y_2) : y_1 - y_2 \in W\}\}, \end{aligned}$$

i.e. \mathcal{N} is the uniformity on Y generating τ_Y . Using these notations, apply 5.3.2 and the principles of idealization and transfer to infer

$$\begin{aligned} &(\forall N \in \mathcal{N}) N(\mu(\mathcal{A})) \supset \mu(\mathcal{B}) \\ \Leftrightarrow &(\forall N \in \mathcal{N})(\forall b \in \mu(\mathcal{B}))(\exists a \in \mu(\mathcal{A}))(b \in N(a)) \\ \Leftrightarrow &(\forall N \in \mathcal{N})(\forall^{\text{st}} A \in \mathcal{A})(\exists^{\text{st}} B \in \mathcal{B})(\forall b \in B)(\exists a \in A)(b \in N(a)) \\ \Leftrightarrow &(\forall^{\text{st}} A \in \mathcal{A})(\forall N \in \mathcal{N})(\exists^{\text{st}} B \in \mathcal{B})(B \subset N(A)) \\ \Leftrightarrow &(\forall^{\text{st}} A \in \mathcal{A})(\exists^{\text{st}} B \in \mathcal{B})(\forall N \in \mathcal{N})(B \subset N(A)) \\ \Leftrightarrow &(\forall^{\text{st}} A \in \mathcal{A})(\exists^{\text{st}} B \in \mathcal{B})(B \subset \text{cl}(A)) \\ \Leftrightarrow &(\forall A \in \mathcal{A})(\exists B \in \mathcal{B})(B \subset \text{cl}(A)), \end{aligned}$$

which completes the proof. ▷

5.5.3. Theorem. *The following hold:*

- (1) *If T , F and x' satisfy condition (ρ) and T is a continuous mapping from (X, τ_X) to (Y, τ_Y) then*

$$\begin{aligned} T(\text{Cl}_\Lambda(F, x')) &\subset \text{Cl}_\Lambda(T(F), Tx'), \\ T(\text{In}_\Lambda(F, x')) &\subset \text{In}_\Lambda(T(F), Tx'); \end{aligned}$$

if, moreover, T is an open mapping from (X, τ_X) to (Y, τ_Y) then

$$T(\text{Ha}_\Lambda(F, x')) \subset \text{Ha}_\Lambda(T(F), T(x'));$$

- (2) *If τ_Y is a vector topology, T , F , and x' satisfy condition $(\bar{\rho})$, while $T : (X, \tau_X) \rightarrow (Y, \tau_Y)$ is continuous then*

$$T(\text{Cl}_\Lambda(F, x')) \subset \text{Cl}_\Lambda(T(F), Tx').$$

◁ (1): By way of example, we will demonstrate the second inclusion. To this end, take $h' \in \text{In}_\Lambda(F, x')$. Given $\alpha \in \Lambda$, choose $h \approx_{\tau_X} h'$ so that $x + \alpha h \in F$ for all $x \approx_{\sigma_X} x', x \in F$. Obviously, $Th \approx_{\sigma_X} Th'$ and $Tx + \alpha Th \in T(F)$. Applying condition (ρ) , conclude that $Th' \in \text{In}_\Lambda(T(F), Tx')$.

Assume now that T is open, i.e.

$$T(\mu(\tau_X)) \supset \mu(\tau_Y)$$

by 5.5.2 (1). Together with the continuity of T , this implies that the above monads coincide. If now $y \in T(F)$, $y \approx_{\sigma_Y} Tx'$, then by condition (ρ) we have $y = Tx$, where $x \in F$ and $x \approx_{\sigma_Y} x'$. Moreover, given $z \approx_{\tau_Y} Th'$, we may find $h \approx_{\tau_X} h'$ with $z = Th$. Therefore, for all $\alpha \in \Lambda$ we have $x + \alpha h \in F$, i.e. $y + \alpha z = Tx + \alpha Th \in T(F)$ as soon as a standard h' is such that $h' \in \text{Ha}_\Lambda(F, x')$.

(2): Take an infinitesimal $\alpha \in \Lambda$ and a standard element $h' \in \text{Cl}_\alpha(F, x')$. Let W be some infinitesimal neighborhood of the origin in τ_Y . Then, by hypothesis, αW is also a neighborhood of the origin. By condition $(\bar{\rho})$, given $y \approx_{\sigma_Y} Tx'$, $y \in T(F)$, we find $x \in \mu(\sigma_X(x')) \cap F$ satisfying $y = Tx + \alpha\omega$ and $\omega \approx_{\tau_Y} 0$. Since h' belongs to the Clarke cone, there is some h'' such that $h'' \approx_{\tau_Y} h'$ and $x + \alpha h'' \in F$. Hence, $y + \alpha(Th'' - w) = y - \alpha w + \alpha Th'' = T(x + \alpha h'') \in T(F)$. Thus, $Th'' - w \in Th' + \mu(\tau_Y) - w \in Th' + \mu(\tau_Y) + \mu(\tau_Y) = Th' + \mu(\tau_Y)$. This yields $Th' \in \text{Cl}_\alpha(T(F), Tx')$. ▷

5.5.4. We now consider vector spaces X, Y , and Z furnished with topologies $\sigma_X, \tau_X, \sigma_Y, \tau_Y$, and σ_Z, τ_Z , respectively. Let $F \subset X \times Y$ and $G \subset X \times Z$ be two correspondences, and let a point $d' := (x', y', z') \in X \times Y \times Z$ meet the conditions $a' := (x', y') \in F$ and $b' := (y', z') \in G$. Put $H := X \times G \cap F \times Z$, $c' := (x', z')$. It is worth recalling that $G \circ F = \text{Pr}_{X \times Z} H$, where $\text{Pr}_{X \times Z}$ is the natural projection to $X \times Z$ along Y . We also introduce the following abbreviations:

$$\begin{aligned} \sigma_1 &:= \sigma_X \times \sigma_Y; & \sigma_2 &:= \sigma_Y \times \sigma_Z; & \sigma &:= \sigma_X \times \sigma_Z; & \bar{\sigma} &:= \sigma_X \times \sigma_Y \times \sigma_Z; \\ \tau_1 &:= \tau_X \times \tau_Y; & \tau_2 &:= \tau_Y \times \tau_Z; & \tau &:= \tau_X \times \tau_Z; & \bar{\tau} &:= \tau_X \times \tau_Y \times \tau_Z. \end{aligned}$$

It is worth recalling that $\text{Pr}_{X \times Z}$ is a continuous and open operator (on assuming “lettersame” topologies). We still distinguish some set Λ of infinitesimals. We also need the next property of monads:

5.5.5. *The monad of a composite is the composite of monads.*

◁ Let \mathcal{A} be a filter on $X \times Y$, and let \mathcal{B} be a filter on $Y \times Z$. Put

$$\mathcal{B} \circ \mathcal{A} := \text{fil}\{B \circ A : A \in \mathcal{A}, B \in \mathcal{B}\},$$

where we may assume all $B \circ A$ nonempty. Clearly,

$$B \circ A = \text{Pr}_{X \times Z}(A \times Z \cap X \times B).$$

Therefore, $\mathcal{B} \circ \mathcal{A}$ is the image $\text{Pr}_{X \times Z}(\mathcal{C})$, where $\mathcal{C} := \mathcal{C}_1 \vee \mathcal{C}_2$, $\mathcal{C}_1 := \mathcal{A} \times \{Z\}$, and $\mathcal{C}_2 := \{X\} \times \mathcal{B}$. Since the monad of a product is the product of monads, the monad of the least upper bound of filters is the intersection of their monads, and the monad of the image of a filter coincides with the image of its monad, we come to the relation

$$\mu(\mathcal{B} \circ \mathcal{A}) = \text{Pr}_{X \times Z}(\mu(\mathcal{A}) \times Z \cap X \times \mu(\mathcal{B})) = \mu(\mathcal{B}) \circ \mu(\mathcal{A}),$$

which was required. \triangleright

5.5.6. *The following are equivalent:*

- (1) Condition (ρ) is fulfilled for $\text{Pr}_{X \times Z}$, H , and c' ;
- (2) $G \circ F \cap \mu(\sigma(c')) = G \cap \mu(\sigma_2(b')) \circ F \cap \mu(\sigma_1(a'))$;
- (3) $(\forall V \in \sigma_Y(y'))(\exists U \in \sigma_X(x'))(\exists W \in \sigma_Z(z'))$

$$G \circ F \cap U \times W \subset G \circ I_V \circ F,$$

with I_V standing as usual for the identity relation on V .

\triangleleft Using 5.3.2, rewrite (3) as

$$\begin{aligned} & (\forall V \in \sigma_Y(y'))(\exists O \in \sigma(c'))(\forall (x, z) \in O, (x, z) \in G \circ F) \\ & (\exists y \in V)(x, y) \in F \wedge (y, z) \in G \leftrightarrow (\forall (x, z) \approx_{\sigma(c')} (x, z) \in G \circ F) \\ & (\exists y \approx_{\sigma_Y} y')(x, y) \in F \wedge (y, z) \in G \leftrightarrow \mu(\sigma(c')) \cap G \circ F \\ & \quad \subset \mu(\sigma_2(b')) \cap G \circ \mu(\sigma_1(a')) \cap F. \end{aligned}$$

It remains to observe that

$$\begin{aligned} & \text{Pr}_{X \times Z}(\mu(\bar{\sigma}(d')) \cap H) \\ & = \{(x, z) \in G \circ F : x \approx_{\sigma_X} x' \wedge z \approx_{\sigma_Z} z' \wedge (\exists y \approx_{\sigma_Y} y')(x, y) \in F \wedge (y, z) \in G\} \\ & = \mu(\sigma_2(b')) \cap G \circ \mu(\sigma_1(a')) \cap F, \end{aligned}$$

which completes the proof. \triangleright

5.5.7. *The following are equivalent:*

- (1) Condition $(\bar{\rho})$ is fulfilled for $\text{Pr}_{X \times Z}$, H , and c' ;
- (2) $(\forall W \in \mathcal{N}_\tau) \mu(\sigma_2(b')) \cap G \circ \mu(\sigma_1(a')) \cap F + W \supset \mu(\sigma(c')) \cap G \circ F$;
- (3) $(\forall V \in \sigma_2(b'))(\forall U \in \sigma_1(a'))(\exists W \in \sigma(c'))$
 $W \cap G \circ F \subset \text{cl}_\tau(V \cap G \circ U \cap F)$;
- (4) $(\forall U \in \sigma_X(x'))(\forall V \in \sigma_Y(y'))(\forall W \in \sigma_Z(z'))(\exists V \in \sigma(c'))$
 $O \cap G \circ F \subset \text{cl}_\tau(G \circ I_V \circ F \cap U \times W)$;
- (5) If $\tau \geq \sigma$ then $(\forall V \in \sigma_Y(y'))(\exists U \in \sigma_X(x'))(\exists W \in \sigma_Z(z'))$
 $G \circ F \cap U \times W \subset \text{cl}_\tau(G \circ I_V \circ F)$

(in this event we say that condition $(\bar{\rho}c)$ is fulfilled for the point $d' := (x', y', z')$).

◁ From 5.5.2 (3) and the proof of 5.5.2 (3) we directly infer that (1)↔(2)↔(3). To prove the equivalence (3) ↔ (4), it suffices to observe that

$$\begin{aligned} & (V \times W) \cap G \circ (U \times V) \cap F \\ &= \{(x, z) \in X \times Z : x \in U \wedge z \in W \wedge (\exists y \in V)(x, y) \in F \wedge (y, z) \in G\} \\ &= G \circ I_V \circ F \cap U \times W \end{aligned}$$

for all $U \subset X$, $V \subset Y$, and $W \subset Z$. We are left with establishing (4) ↔ (5).

To this end, note that (4) → (5) is obvious since (5) is a specialization of (4) with $U := X$ and $W := Z$. To validate (5) → (4), take $V \in \sigma_Y(y')$ and select an open neighborhood $C \in \sigma(c')$ such that $G \circ F \cap C \subset \text{cl}_\tau(A)$, where $A := G \circ I_V \circ F$. Given open sets $U \in \sigma_X(x')$ and $W \in \sigma_Z(z')$, put $B := U \times W$ and $O := B \cap C$. Obviously, $G \circ F \cap O \subset (\text{cl}_\tau(A)) \cap B$. Assuming the standard environment and granted $a \in (\text{cl}_\tau(A)) \cap B$, find a point $a' \in A$ satisfying $a' \approx_\tau a$. Clearly, $a' \approx_\sigma a$, since $\mu(\tau) \subset \mu(\sigma)$ by hypothesis. The set B is σ -open and so $a' \in B$, i.e. $a' \in A \cap B$ and $a \in \text{cl}_\tau(A \cap B)$. Finally, $G \circ F \cap O \subset \text{cl}_\tau(A \cap B)$, as required. ▷

5.5.8. *The following hold:*

- (1) $\text{Ha}_\Lambda(H, d') \supset X \times \text{Ha}_\Lambda(G, b') \cap \text{Ha}_\Lambda(F, a') \times Z$;
 - (2) $\text{R}_\Lambda^2(H, d') \supset X \times \text{R}_\Lambda^1(G, b') \cap \text{R}_\Lambda^2(F, a') \times Z$;
 - (3) $\text{Cl}_\Lambda(H, d') \supset X \times \text{Q}_\Lambda^1(G, b') \cap \text{Cl}_\Lambda(F, a') \times Z$;
 - (4) $\text{Cl}_\Lambda(H, d') \supset X \times \text{Cl}(G, b') \cap \text{Q}_\Lambda^2(F, a') \times Z$;
 - (5) $\text{Cl}^2(H, d') \supset X \times P^2(G, b') \cap \text{S}^2(F, a') \times Z$,
- where $\text{Cl}^2(H, d')$ is defined as

$$\begin{aligned} & \text{Cl}^2(H, d') \\ &:= * \{(s', t', r') \in X \times Y \times Z : (\forall d \approx_{\bar{\sigma}} d', d \in H) \\ & (\forall \alpha \in \mu(\mathbb{R}_+)) (\exists s \approx_{\tau_X} s') (\forall t \approx_{\tau_Y} t') (\exists r \approx_{\tau_Z} z') \\ & (d + \alpha(s, t, r) \in H)\}. \end{aligned}$$

◁ We will check only (1) and (5), since the remaining claims are provable by analogy.

(1): Assume that (s', t', r') is a standard element belonging to the right side. Take $d \approx_{\bar{\sigma}} d'$ and $\alpha \in \Lambda$, with $d := (x, y, z) \in H$. Clearly, $a := (x, y) \in F$ and $a \approx_{\sigma_1} a'$, while $b := (y, z) \in G$ and $b \approx_{\sigma_2} b'$. Therefore, $a + \alpha(s, t) \in F$ and $b + \alpha(t, r) \in G$ for all $\alpha \in \Lambda$ and $(s, t, r) \approx_{\bar{\tau}} (s', t', r')$. Hence,

$$\begin{aligned} d + \alpha(s, t, r) &= (a + \alpha(s, t), z + \alpha r) \in F \times Z, \\ d + \alpha(s, t, r) &= (x + \alpha s, b + \alpha(t, r)) \in X \times G, \end{aligned}$$

i.e. $(s', t', r') \in \text{Ha}_\Lambda(H, d')$.

(5): Take a standard element (s', t', r') from the right of (4). By definition, we may find s such that $s \approx_{\tau_X} s'$ and to each $t \approx_{\tau_Y} t'$ there is some $r \approx_{\tau_Z} r'$ satisfying $a + \alpha(s, t) \in F$ and $b + \alpha(t, r) \in G$ for all $a \approx_{\sigma_1} a'$ and $b \approx_{\sigma_2} b'$. Obviously, $d + \alpha(s, t, r) \in H$ whenever $b \approx_{\bar{\sigma}} d'$ and $d \in H$. \triangleright

5.5.9. It should be emphasized that the mechanism of “leapfrogging” in 5.5.8 can be modified so as to meet the aims of research. These aims usually include convenient approximation to composites. We note in passing that the scheme is fruitful resting on the method of general position [263, 279]. By way of illustration, we state the following typical result.

5.5.10. Theorem. *Let τ be a vector topology with $\tau \geq \sigma$. Let $F \subset X \times Y$ and $G \subset Y \times Z$ be correspondences such that $\text{Ha}(F, a') \neq \emptyset$. Assume further that the cones $Q^2(F, a') \times Z$ and $X \times \text{Cl}(G, b')$ are in general position (relative to the topology $\bar{\tau}$). Then*

$$\text{Cl}(G \circ F, c') \supset \text{Cl}(G, b') \circ \text{Cl}(F, a'),$$

whenever condition $(\bar{\rho}c)$ is fulfilled at d' .

\triangleleft The proof proceeds along the lines of [263] (compare 5.3.13) and consists in verifying the conditions that guarantee the validity for the following estimating

$$\begin{aligned} \text{Cl}(G \circ F, c') &= \text{Cl}(\text{Pr}_{X \times Z} H, \text{Pr}_{X \times Z} d') \supset \text{cl}_\tau(\text{Pr}_{X \times Z} \text{Cl}(H, d')) \\ &\supset \text{Pr}_{X \times Z} \text{cl}_{\bar{\tau}}(X \times \text{Cl}(G, b') \cap Q^2(F, a') \times Z) \\ &= \text{Pr}_{X \times Z}(\text{cl}_{\bar{\tau}}(X \times \text{Cl}(G, b')) \cap \text{cl}_{\bar{\tau}}(Q^2(F, a') \times Z)) \\ &= \text{Pr}_{X \times Z}(X \times \text{Cl}(G, b') \cap \text{Cl}(F, a') \times Z) = \text{Cl}(G, b') \circ \text{Cl}(F, a'). \end{aligned}$$

The proof is complete. \triangleright

5.6. Infinitesimal Subdifferentials

Optimization theory pays attention to the problem of how the accuracy of constraints and optimality criteria influences solutions and values in numerical calculation. One of the qualitative approach to this problem is reflected in the so-called convex ε -programming which provides tools for estimating approximations to an optimum “by functional,” i.e. by the values of the objective function. The technique of ε -programming is rather specific and artificially complicated in a sense in regard to its recommendations for recalculating accuracy as ε varies. These recommendations in a form of ε -subdifferential calculus are in an outright contradiction with the common practice. The latter rests on belief that we are close to a “practical optimum” whenever we satisfy the complementary slackness conditions or their versions corresponding to the classical case of $\varepsilon = 0$ with “practical accuracy.” This

results in a glaring discrepancy if not an abyss between the theory and practice of optimization.

In this section we outline an approach to bridging the gap within the radical stance of infinitesimal analysis. The basic tool is the concept of infinitesimal optimal solution which is a feasible point at which the objective function is infinitely close to the ideal, not necessarily attained value of the program under study. The infinitesimal optimum appears so an acceptable candidate for the role of a “practical” optimum, since no “assignable” procedures can differ such an optimum from the conventional, “theoretical” optimum. We derive the main formulas for calculating the so-called infinitesimal subdifferentials that reflect the new concept of infinitesimal optimality. It is worth emphasizing that the resulting formulas for external sets coincide in form with their classical analogs in standard convex analysis. Moreover, the new criteria for infinitesimal optimality actually involve the approximate complementary slackness.

5.6.1. Considering a vector space X , we let E^\bullet be an ordered vector space extended with a greatest element $+\infty$. Assume given a convex operator $f : X \rightarrow F^\bullet$ and a point \bar{x} in the effective domain $\text{dom}(f) := \{x \in X : f(x) < +\infty\}$ of F . Given $\varepsilon \geq 0$ in the positive cone E_+ of E , by the ε -subdifferential of f at \bar{x} we mean the set

$$\partial^\varepsilon f(\bar{x}) := \{T \in L(X, E) : (\forall x \in X)(Tx - Fx \leq T\bar{x} - f\bar{x} + \varepsilon)\},$$

with $L(X, E)$ standing as usual for the space of linear operators from X to E .

5.6.2. Distinguish some downward-filtered subset \mathcal{E} of E composed of positive elements. Assuming E and \mathcal{E} standard, define the monad $\mu(\mathcal{E})$ of \mathcal{E} as $\mu(\mathcal{E}) := \bigcap \{[0, \varepsilon] : \varepsilon \in {}^\circ\mathcal{E}\}$. The members of $\mu(\mathcal{E})$ are *positive infinitesimals* with respect to \mathcal{E} . As usual, ${}^\circ\mathcal{E}$ denotes the external set of standard members of E , the *standard part* of \mathcal{E} . Observe that, adopting the canons of this field of research, we follow the radical stance of nonstandard set theory.

Without further specification we assume in the sequel that E is *Kantorovich space* or a *K-space* also known as a Dedekind complete vector lattice. We will assume that the monad $\mu(\mathcal{E})$ is an external cone over ${}^\circ\mathbb{R}$ and, moreover, $\mu(\mathcal{E}) \cap {}^\circ E = 0$. (In application, \mathcal{E} is usually the filter of order-units of E .) The relation of *infinite proximity* or *infinite closeness* between the members of E is introduced as follows:

$$e_1 \approx e_2 \leftrightarrow e_1 - e_2 \in \mu(\mathcal{E}) \wedge e_2 - e_1 \in \mu(\mathcal{E}).$$

5.6.3. *The equality holds:*

$$\bigcap_{\varepsilon \in {}^\circ\mathcal{E}} \partial_\varepsilon f(\bar{x}) = \bigcup_{\varepsilon \in \mu(\mathcal{E})} \partial_\varepsilon f(\bar{x}).$$

◁ Given $T \in L(X, E)$, proceed successively as follows

$$\begin{aligned}
T \in \bigcap_{\varepsilon \in {}^\circ\mathcal{E}} \partial_\varepsilon f(\bar{x}) &\leftrightarrow (\forall^{\text{st}} \varepsilon \in \mathcal{E})(\forall x \in X)(Tx - T\bar{x} \leq f(x) - f(\bar{x}) + \varepsilon) \\
&\leftrightarrow (\forall^{\text{st}} \varepsilon \in \mathcal{E}) f^*(T) := \sup_{x \in \text{dom}(f)} (Tx - f(x)) \leq T\bar{x} - f(\bar{x}) + \varepsilon \\
&\leftrightarrow (\forall^{\text{st}} \varepsilon \in \mathcal{E}) 0 \leq f^*(T) - (T\bar{x} - f(\bar{x})) \leq -\varepsilon \\
&\leftrightarrow f^*(T) - (T\bar{x} - f(\bar{x})) \approx 0 \\
&\leftrightarrow (\exists \varepsilon \in E^+) \varepsilon \approx 0 \wedge f^*(T) = T\bar{x} - f(\bar{x}) + \varepsilon \\
&\leftrightarrow T \in \bigcup_{\varepsilon \in \mu(\mathcal{E})} \partial_\varepsilon f(\bar{x}),
\end{aligned}$$

The proof is complete. ▷

5.6.4. The external set on both sides of 5.6.3 is the *infinitesimal subdifferential* of f at \bar{x} . We denote this set by $Df(\bar{x})$. The elements of $Df(\bar{x})$ are *infinitesimal subgradients* of f at \bar{x} . We abstain from explicitly indicating the set \mathcal{E} since this leads to no confusion.

5.6.5. Assume the standard environment; i.e., let the parameters X, f, \bar{x} , etc. be standard. Then the standardization of the infinitesimal subdifferential of f at \bar{x} coincides with the (zero) subdifferential of f at \bar{x} , i.e.

$${}^*Df(\bar{x}) = \partial f(\bar{x}).$$

◁ Given a standard $T \in L(X, E)$, infer by transfer that

$$\begin{aligned}
T \in {}^*Df(\bar{x}) &\leftrightarrow T \in Df(\bar{x}) \\
&\leftrightarrow (\forall^{\text{st}} \varepsilon \in \mathcal{E})(\forall x \in X)(Tx - T\bar{x} \leq f(x) - f(\bar{x}) + \varepsilon) \\
&\leftrightarrow (\forall \varepsilon \in \mathcal{E})(\forall x \in X)(Tx - T\bar{x} \leq f(x) - f(\bar{x}) + \varepsilon) \\
&\leftrightarrow T \in \partial f(\bar{x}).
\end{aligned}$$

The above uses the equality $\inf \mathcal{E} = 0$ which ensues from the assumption that $\mu(\mathcal{E}) \cap {}^\circ E = 0$. ▷

5.6.6. Let F be a standard Kantorovich space and let $g : E \rightarrow F^\bullet$ be an increasing convex operator. If $E \times \text{epi}(g)$ and $\text{epi}(f) \times G$ are in general position then

$$D(g \circ f)(\bar{x}) = \bigcup_{T \in Dg(f(\bar{x}))} D(T \circ f)(\bar{x}).$$

If, moreover, the parameters are standard (with possible exception of \bar{x}) then

$${}^\circ D(g \circ f)(\bar{x}) = \bigcup_{T \in {}^\circ Dg(f(\bar{x}))} {}^\circ D(T \circ f)(\bar{x}).$$

◁ Observe that by assumption the monad $\mu(\mathcal{E})$ is an external normal subgroup in F , i.e.,

$$\begin{aligned} \varepsilon \in \mu(\mathcal{E}) &\rightarrow [0, \varepsilon] \subset \mu(\mathcal{E}), \\ \mu(\mathcal{E}) + \mu(\mathcal{E}) &\subset \mu(\mathcal{E}). \end{aligned}$$

Using 4.6.3 and the rules for calculating ε -subdifferentials (see 4.2.11 (2)), proceed successively as follows

$$\begin{aligned} D(g \circ f)(\bar{x}) &= \bigcup_{\varepsilon \in \mu(\mathcal{E})} \partial_\varepsilon(g \circ f)(\bar{x}) \\ &= \bigcup_{\varepsilon \in \mu(\mathcal{E})} \bigcup_{\substack{\varepsilon_1 + \varepsilon_2 = \varepsilon \\ \varepsilon_1 \geq 0, \varepsilon_2 \geq 0}} \bigcup_{T \in \partial_{\varepsilon_1} g(f(\bar{x}))} \partial_{\varepsilon_2}(T \circ f)(\bar{x}) \\ &= \bigcup_{\substack{\varepsilon_1 \geq 0, \varepsilon_2 \geq 0 \\ \varepsilon_1 \approx 0, \varepsilon_2 \approx 0}} \bigcup_{T \in \partial_{\varepsilon_1} g(f(\bar{x}))} \partial_{\varepsilon_2}(T \circ f)(\bar{x}) \\ &= \bigcup_{\varepsilon_1 \geq 0, \varepsilon_1 \approx 0} \bigcup_{T \in \partial_{\varepsilon_1} g(f(\bar{x}))} \bigcup_{\varepsilon_2 \geq 0, \varepsilon_2 \approx 0} \partial_{\varepsilon_2}(T \circ f)(\bar{x}) \\ &= \bigcup_{\varepsilon_1 \geq 0, \varepsilon_1 \approx 0} \bigcup_{T \in \partial_{\varepsilon_1} g(f(\bar{x}))} D(T \circ f)(\bar{x}). \end{aligned}$$

Assume the standard environment and take $S \in {}^\circ D(g \circ f)(\bar{x})$. Then

$$(g \circ f)^*(S) = \sup_{x \in \text{dom}(g \circ f)} (Sx - g \circ f(x)) \leq S\bar{x} - g(f(\bar{x})) + \varepsilon$$

for some infinitesimal ε . By transfer and the change-of-variable formula for the Young–Fenchel transform, there is a standard operator $T \in {}^\circ L(E, F)$ such that T is positive, i.e. $T \in L^+(E, F)$, and moreover

$$(g \circ f)^*(S) = (T \circ f)^*(S) + g^*(T).$$

This implies

$$\begin{aligned}
\varepsilon &\geq \sup_{x \in \text{dom}(f)} (Sx - T \circ f(x)) + \sup_{e \in \text{dom}(g)} (Te - g(e)) - S\bar{x} + g(f(\bar{x})) \\
&= \sup_{x \in \text{dom}(f)} (Sx - S\bar{x} - (T \circ f(x) - T \circ f(\bar{x}))) \\
&\quad + \sup_{e \in \text{dom}(g)} (Te - T \circ f(\bar{x}) - (g(e) - g(f(\bar{x}))))).
\end{aligned}$$

Put

$$\begin{aligned}
\varepsilon_1 &:= \sup_{e \in \text{dom}(g)} (Te - T \circ f(\bar{x}) - (g(e) - g(f(\bar{x}))))), \\
\varepsilon_2 &:= \sup_{x \in \text{dom}(f)} (Sx - S\bar{x} - (T \circ f(x) - T \circ f(\bar{x}))).
\end{aligned}$$

Clearly, $S \in \partial_{\varepsilon_2}(T \circ f)(\bar{x})$, i.e., $S \in {}^\circ D(T \circ f)(\bar{x})$; and $T \in \partial_{\varepsilon_1}(g \circ f)(\bar{x})$, i.e., $T \in {}^\circ Dg(f(\bar{x}))$ since $\varepsilon_1 \approx 0$ and $\varepsilon_2 \approx 0$. \triangleright

5.6.7. Let $f_1, \dots, f_n : X \rightarrow E^\bullet$ be convex operators and let n be a standard number. If f_1, \dots, f_n are in general position then

$$D(f_1 + \dots + f_n)(\bar{x}) = Df_1(\bar{x}) + \dots + Df_n(\bar{x})$$

for $\bar{x} \in \text{dom}(f_1) \cap \dots \cap \text{dom}(f_n)$.

\triangleleft The proof consists in applying 5.6.3 and the rules for ε -subdifferentiation of a sum on observing that the sum of standardly many infinitesimal summands is infinitesimal too. \triangleright

5.6.8. Let $f_1, \dots, f_n : X \rightarrow E^\bullet$ be convex operators and let n be a standard number. Assume that f_1, \dots, f_n are in general position, E is a vector lattice, and $\bar{x} \in \text{dom}(f_1 \vee \dots \vee f_n)$. If F is a standard Kantorovich space and $T \in L^+(E, F)$ is a positive linear operator then $S \in L(X, F)$ is an infinitesimal subgradient of the operator $T \circ (f_1 \vee \dots \vee f_n)$ at \bar{x} if and only if the following system of conditions is compatible:

$$\begin{aligned}
T &= \sum_{k=1}^n T_k; \quad T_k \in L^+(E, F) \quad (k := 1, \dots, n); \\
\sum_{k=1}^n T_k \bar{x} &\approx T(f_1(\bar{x}) \vee \dots \vee f_n(\bar{x})); \quad S \in \sum_{k=1}^n D(T_k \circ f_k)(\bar{x}).
\end{aligned}$$

\triangleleft Define the operators

$$(f_1, \dots, f_n) : X \rightarrow (E^n)^\bullet, \quad (f_1, \dots, f_n)(x) := (f_1(x), \dots, f_n(x));$$

$$\varkappa : E^n \rightarrow E, \quad \varkappa(e_1, \dots, e_n) := e_1 \vee \dots \vee e_n.$$

Then

$$T \circ f_1 \vee \dots \vee f_n = T \circ \varkappa \circ (f_1, \dots, f_n).$$

Since $T \circ \varkappa$ is a sublinear operator, we are done on appealing to 5.6.5. \triangleright

5.6.9. Let X be a vector space, let E be some Kantorovich space, and let \mathcal{A} be a weakly order bounded set in $L(X, E)$. Look at the regular convex operator

$$f = \varepsilon_{\mathcal{A}} \circ \langle \mathcal{A} \rangle^e;$$

where, as usual, $\varepsilon_{\mathcal{A}}$ is the canonical sublinear operator

$$\varepsilon_{\mathcal{A}} : l_\infty(\mathcal{A}, E) \rightarrow E, \quad \varepsilon_{\mathcal{A}}(f) := \sup f(\mathcal{A})$$

and, for $e \in l_\infty(\mathcal{A}, E)$, the affine operator $\langle \mathcal{A} \rangle^e$ acts by the rule

$$\langle \mathcal{A} \rangle^e x := \langle \mathcal{A} \rangle x + e, \quad \langle \mathcal{A} \rangle x : T \in \mathcal{A} \rightarrow Tx.$$

5.6.10. Let $g : E \rightarrow F^\bullet$ be an increasing convex operator acting to a standard Kantorovich space F . Assume that the image $f(X)$ contains an algebraically internal point of $\text{dom}(g)$; and let \bar{x} be an element in X such that $f(\bar{x}) \in \text{dom}(g)$. Then

$$D(g \circ f)(\bar{x})$$

$$= \{T \circ \langle \mathcal{A} \rangle : T \circ \Delta_{\mathcal{A}} \in Dg(f(\bar{x})), T \geq 0, T \circ \Delta_{\mathcal{A}} f(\bar{x}) \approx T \circ \langle \mathcal{A} \rangle^e \bar{x}\}.$$

\triangleleft If $S \in D(g \circ f)(\bar{x})$ then, by 5.6.3, $S \in \partial_\varepsilon(g \circ f)(\bar{x})$ for some $\varepsilon \approx 0$. It remains to appeal to the respective rule of ε -differentiation. If $T \geq 0$, $T \circ \Delta_{\mathcal{A}} \in Dg(f(\bar{x}))$, and $T \circ \Delta_{\mathcal{A}} f(\bar{x}) \approx T \circ \langle \mathcal{A} \rangle^e \bar{x}$ then, for some $\varepsilon \approx 0$, $T \circ \Delta_{\mathcal{A}} \in \partial_\varepsilon g(f(\bar{x}))$. Put $\delta := T \circ \Delta_{\mathcal{A}} f(\bar{x}) - T \circ \langle \mathcal{A} \rangle^e \bar{x}$. Then $\delta \geq 0$ and $\delta \approx 0$ by hypothesis. Hence, $T \circ \langle \mathcal{A} \rangle \in \partial_{\varepsilon+\delta}(g \circ f)(\bar{x})$. It remains to observe that $\varepsilon + \delta \approx 0$. \triangleright

5.6.11. Under the assumptions of 5.6.10, let g be a sublinear Maharam operator [268]. Then

$$D(g \circ f)(\bar{x}) = \bigcup_{T \in Dg(f(\bar{x}))} \bigcup_{\delta \geq 0, T\delta \approx 0} T(\partial_\delta f(\bar{x})).$$

◁ By 5.6.5, we may assume that $g := T$. If $T\delta \approx 0$ and

$$Cx - C\bar{x} \leq f(x) - f(\bar{x}) + \delta$$

for all $x \in X$ then it is clear that

$$TC \in \partial_{T\delta}(T \circ f)(\bar{x}) \subset D(T \circ f)(\bar{x}).$$

To complete the proof, take $S \in D(T \circ f)(\bar{x})$. By 4.6.3, there is an infinitesimal ε such that $S \in \partial_\varepsilon(T \circ f)(\bar{x})$. Appealing to the appropriate rule of ε -subdifferentiation, find some $\delta \geq 0$ and $C \in \partial_\delta f(\bar{x})$ so that $T\delta \leq \varepsilon$ and $S = TC$. This completes the proof. ▷

5.6.12. Let A be some set and let $(f_\alpha)_{\alpha \in A}$ be a uniformly regular family of convex operators. Then

$$\begin{aligned} D\left(\sum_{\xi \in \Xi} f_\xi\right)(\bar{x}) &= \bigcup_{\substack{\delta \in l_1(\Xi, E) \\ \delta \geq 0, \delta \approx 0}} \sum_{\xi \in \Xi} \partial_{\delta(\xi)} f_\xi(\bar{x}); \\ &D\left(\sup_{\xi \in \Xi} f_\xi\right)(\bar{x}) \\ &= \bigcup \left\{ \sum_{\xi \in \Xi} \alpha_\xi \partial_{\delta(\xi)} f_\xi(\bar{x}) : 0 \leq \alpha_\xi \leq \mathbb{1}_E, \sum_{\xi \in \Xi} \alpha_\xi = \mathbb{1}_E, \right. \\ &\quad \left. \sum_{\xi \in \Xi} \alpha_\xi f_\xi(\bar{x}) \approx \sup_{\xi \in \Xi} f_\xi(\bar{x}), \sum_{\xi \in \Xi} \alpha_\xi \delta(\xi) \approx 0 \right\}. \end{aligned}$$

◁ The claim is immediate from 5.6.11 on recalling the rules for disintegration (see [279]). ▷

5.6.13. It is worth observing that the formulas of 5.6.7–5.6.12 admit refinements in analogy with 5.6.6 on assuming the standard environment (with the possible exception of \bar{x}). We also emphasize that, proceeding along the above lines, we may derive the whole spectrum of all formulas of subdifferential calculus (rules for infimal convolutions, level sets, etc.).

5.6.14. Assume as before that $f : X \rightarrow E^\bullet$ is a convex operator acting to a standard Kantorovich space E , and let $\mathcal{X} := \mathcal{X}(\cdot)$ be a *generalized point* in $\text{dom}(f)$, i.e. a net of elements in $\text{dom}(f)$. An operator T in $L(X, E)$ is an *infinitesimal gradient* of f at \mathcal{X} provided that

$$f^*(T) \leq \liminf (T\mathcal{X} - f(\mathcal{X})) + \varepsilon$$

for some positive infinitesimal ε . (Of course, we presume the rule $T\mathcal{X} := T \circ \mathcal{X}$ here.) Assuming the standard environment, we see that an infinitesimal subgradient is merely a support operator at a generalized point.

We agree to denote by the symbol $Df(\mathcal{X})$ the set of infinitesimal subgradients of f at \mathcal{X} . This set is naturally called the *infinitesimal subdifferential* of f at \mathcal{X} .

We will now derive two main rules for subdifferentiation at a generalized point which are of profound interest since no exact formulas are available for the respective ε -subdifferentials.

5.6.15. *Let f_1, \dots, f_n be a collection of standard convex operators in general position and let a generalized point \mathcal{X} belong to $\text{dom}(f_1) \cap \dots \cap \text{dom}(f_n)$. Then*

$$D(f_1 + \dots + f_n)(\mathcal{X}) = Df_1(\mathcal{X}) + \dots + Df_n(\mathcal{X}).$$

◁ Take $T_k \in Df_k(\mathcal{X})$ for $k := 1, \dots, n$, i.e.,

$$f_k^*(T_k) \leq \liminf(T_k\mathcal{X} - f_k(\mathcal{X})) + \varepsilon_k$$

for some infinitesimal $\varepsilon_1, \dots, \varepsilon_n$. In this case,

$$\begin{aligned} (f_1 + \dots + f_n)^*(T_1 + \dots + T_n) &\leq \sum_{k=1}^n f_k^*(T_k) \\ &\leq \sum_{k=1}^n (\liminf(T_k\mathcal{X} - f_k(\mathcal{X})) + \varepsilon_k) \\ &\leq \liminf \sum_{k=1}^n (T_k\mathcal{X} - f_k(\mathcal{X})) + \sum_{k=1}^n \varepsilon_k \end{aligned}$$

by the usual properties of the Young–Fenchel transform and lower limit. It remains to observe that $\varepsilon_1 + \dots + \varepsilon_n \approx 0$ and so the inclusion \supset holds in the formula under proof.

To verify the reverse inclusion, reduce everything to the case $n = 2$ and take $T \in D(f_1 + f_2)(\mathcal{X})$. Then

$$\begin{aligned} (f_1 + f_2)^*(T) &= f_1^*(T_1) + f_2^*(T_2), \\ f_1^*(T_1) + f_2^*(T_2) - \liminf(T\mathcal{X} - (f_1 + f_2)(\mathcal{X})) &\leq \varepsilon \end{aligned}$$

for some $\varepsilon \approx 0$, T_1 , and T_2 such that $T_1 + T_2 = T$. Put

$$\delta_1 := f_1^*(T_1) - \liminf(T_1\mathcal{X} - f_1(\mathcal{X})),$$

$$\delta_2 := f_2^*(T_2) - \liminf(T_2 \mathcal{X} - f_2(\mathcal{X})).$$

For $k := 1, 2$, it is clear that

$$0 \leq \sup_{x \in \text{dom}(f_k)} (T_k x - f_k(x)) - \limsup(T_k \mathcal{X} - f_k(\mathcal{X})) \leq \delta_k.$$

Hence, we are left with verifying that δ_1 and δ_2 are both infinitesimal. To this end, note that

$$\begin{aligned} & \delta_1 + \delta_2 \\ & \leq \varepsilon + \liminf(T \mathcal{X} - (f_1 + f_2)(\mathcal{X})) - \sum_{k=1}^2 \liminf(T_k \mathcal{X} - f_k(\mathcal{X})) \\ & \leq (\varepsilon + \limsup(T_1 \mathcal{X} - f_1(\mathcal{X})) - \liminf(T_1 \mathcal{X} - f_1(\mathcal{X}))) \\ & \quad \wedge (\varepsilon + \limsup(T_2 \mathcal{X} - f_2(\mathcal{X})) - \liminf(T_2 \mathcal{X} - f_2(\mathcal{X}))) \\ & \leq (\varepsilon + f_1^*(T_1) - \liminf(T_1 \mathcal{X} - f_1(\mathcal{X}))) \\ & \quad \wedge (\varepsilon + f_2^*(T_2) - \liminf(T_2 \mathcal{X} - f_2(\mathcal{X}))) \leq \varepsilon + \delta_1 \wedge \delta_2. \end{aligned}$$

Hence, $0 \leq \delta_1 \vee \delta_2 \leq \varepsilon$, which completes the proof. \triangleright

5.6.16. *Let F be a standard Kantorovich space and let $g : E \rightarrow F^\bullet$ be an increasing convex operator. If $X \times \text{epi}(g)$ and $\text{epi}(f) \times F$ are in general position then*

$$D(g \circ f)(\mathcal{X}) = \bigcup_{T \in Dg(f(\mathcal{X}))} D(T \circ f)(\mathcal{X})$$

for a generalized point \mathcal{X} in $\text{dom}(g \circ f)$.

\triangleleft Assume that

$$\begin{aligned} (T \circ f)^*(S) & \leq \liminf(S \mathcal{X} - T \circ f(\mathcal{X})) + \varepsilon_1, \\ g^*(T) & \leq \liminf(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})) + \varepsilon_2 \end{aligned}$$

for some infinitesimals ε_1 and ε_2 . Then

$$\begin{aligned} (g \circ f)^*(S) & \leq (T \circ f)^*(S) + g^*(T) \\ & \leq \liminf(S \mathcal{X} - T \circ f(\mathcal{X})) + \varepsilon_1 + \liminf(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})) + \varepsilon_2 \\ & \leq \liminf(S \mathcal{X} - g \circ f(\mathcal{X})) + \varepsilon_1 + \varepsilon_2. \end{aligned}$$

Consequently, $S \in D(g \circ f)(\mathcal{X})$ and the right side of the formula under study symbolizes the set included into the left side.

To complete the proof, take $S \in D(g \circ f)(\mathcal{X})$. Then there are an infinitesimal ε and an operator T satisfying

$$(g \circ f)^*(S) = (T \circ f)^*(S) + g^*(T) \leq \liminf(S\mathcal{X} - g \circ f(\mathcal{X})) + \varepsilon.$$

Put

$$\begin{aligned} \delta_1 &:= (T \circ f)^*(S) - \liminf(S\mathcal{X} - T \circ f(\mathcal{X})), \\ \delta_2 &:= g^*(T) - \liminf(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})). \end{aligned}$$

Using the properties of upper and lower limits, deduce, first, that

$$\begin{aligned} \delta_1 &\geq (T \circ f)^*(S) - \limsup(S\mathcal{X} - T \circ f(\mathcal{X})) \geq 0, \\ \delta_2 &\geq g^*(T) - \limsup(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})) \geq 0 \end{aligned}$$

and, second, that

$$\begin{aligned} &\delta_1 + \delta_2 \\ &\leq \liminf(S\mathcal{X} - g \circ f(\mathcal{X})) + \varepsilon - \liminf(S\mathcal{X} - T \circ f(\mathcal{X})) \\ &\quad - \liminf(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})) \leq (\limsup(S\mathcal{X} - T \circ f(\mathcal{X})) \\ &\quad - \liminf(S\mathcal{X} - T \circ f(\mathcal{X})) + \varepsilon) \wedge (\limsup(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})) \\ &\quad - \liminf(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})) + \varepsilon) \leq \delta_1 \wedge \delta_2 + \varepsilon, \end{aligned}$$

since we obviously have

$$\begin{aligned} \limsup(T \circ f(\mathcal{X}) - g \circ f(\mathcal{X})) &\leq g^*(T), \\ \limsup(S\mathcal{X} - T \circ f(\mathcal{X})) &\leq (T \circ f)^*(S). \end{aligned}$$

Thus, $0 \leq \delta_1 \vee \delta_2 \leq \varepsilon$, $\delta_1 \approx 0$, and $\delta_2 \approx 0$. This means that $T \in Dg(f(\mathcal{X}))$ and $S \in D(T \circ f)(\mathcal{X})$. \triangleright

5.6.17. We now give an abstraction of the concept of infinitesimal subdifferential which involves the widest spectrum of external possibilities.

Assume as before that F is a convex operator and B is a possibly external subset of $\text{dom}(F)$. Put

$$DF(B) = \bigcap_{\bar{x} \in B} DF(\bar{x}).$$

The external set $DF(B)$ is the *infinitesimal subdifferential* of F along B .

We now take some (generally, external) filterbase \mathcal{B} in the effective domain $\text{dom}(F)$ of a convex operator F . Such a filterbase is sometimes referred to as *generalized point*. Define the infinitesimal subdifferential of F along \mathcal{B} (at the generalized point \mathcal{B}) by the rule

$$DF(\mathcal{B}) := \bigcup_{B \in \mathcal{B}} DF(B).$$

5.6.18. For an operator T in $L(X, E)$, the following are equivalent:

- (1) $T \in DF(\mathcal{B})$;
- (2) $(\exists B \in \mathcal{B})(\forall \bar{x} \in B)(\exists \varepsilon \in \mu(\mathcal{E})) T \in \partial^\varepsilon F(\bar{x})$;
- (3) $(\exists B \in \mathcal{B})(\forall \varepsilon \in {}^\circ\mathcal{E})(\forall \bar{x} \in B) T \in \partial^\varepsilon F(\bar{x})$;
- (4) $(\exists B \in \mathcal{B})(\forall \bar{x} \in B)(\exists \varepsilon \in \mu(\mathcal{E}))(F^*(T) \leq T\bar{x} - F\bar{x} + \varepsilon)$,

with F^* the Young–Fenchel transform of F ;

- (5) $(\exists B \in \mathcal{B})(\forall \bar{x} \in B) \sup_{x \in \text{dom}(F)} ((Tx - T\bar{x}) - (Fx - F\bar{x})) \approx 0$.

◁ Using the definitions, infer

$$\begin{aligned} DF(\mathcal{B}) &= \bigcup_{B \in \mathcal{B}} \bigcap_{\bar{x} \in B} DF(\bar{x}) = \bigcup_{B \in \mathcal{B}} \bigcap_{\bar{x} \in B} \bigcap_{\varepsilon \in {}^\circ\mathcal{E}} \partial^\varepsilon F(\bar{x}) \\ &= \bigcup_{B \in \mathcal{B}} \bigcap_{\varepsilon \in {}^\circ\mathcal{E}} \bigcap_{\bar{x} \in B} \partial^\varepsilon F(\bar{x}), \end{aligned}$$

which amounts to (1) \leftrightarrow (3). Citing the Cauchy principle ensures (2) \leftrightarrow (3). The remaining equivalences follow from the definition of the Young–Fenchel transform. ▷

5.6.19. Let $\mathcal{C} := \{C \subset X : (\exists B \in \mathcal{B}) C \supset B\}$ be the external filter with base \mathcal{B} . Then $DF(\mathcal{C}) = DF(\mathcal{B})$.

◁ It is clear that $\mathcal{C} \supset \mathcal{B}$ and so

$$DF(\mathcal{C}) = \bigcup_{C \in \mathcal{C}} DF(C) \supset \bigcup_{B \in \mathcal{B}} DF(B) = DF(\mathcal{B}).$$

Now, if $T \in DF(\mathcal{C})$ then

$$(\forall \bar{x} \in C) \sup_{x \in \text{dom}(F)} ((Tx - T\bar{x}) - (Fx - F\bar{x})) \approx 0$$

for some C in \mathcal{C} by 5.6.18. The set C includes some member B of the filterbase \mathcal{B} by hypothesis. Appealing to 5.6.18, deduce $T \in DF(B) \subset DF(\mathcal{B})$. ▷

5.6.20. Let \mathcal{B} be an internal filter on X . Assume further that $f : X \rightarrow \mathbb{R}$ is an (everywhere defined) convex function. Given $x^\# \in X^\#$, we then have

$$x^\# \in Df(\mathcal{B}) \leftrightarrow (\exists \varepsilon \in \mu(\mathbb{R}_+)) f^*(x^\#) \leq \liminf (x^\#(\mathcal{B}) - f(\mathcal{B})) + \varepsilon$$

with $\mu(\mathbb{R}_+)$ the set of positive infinitesimals in \mathbb{R} .

◁ To demonstrate the implication to the right, observe that by 5.6.18

$$f^*(x^\#) \leq \inf \{ \langle x \mid x^\# \rangle - f(x) : x \in B \} + \varepsilon$$

for some internal B in \mathcal{B} and every standard $\varepsilon > 0$.

This implies that

$$(\forall \varepsilon \in {}^\circ\mathbb{R}_+) f^*(x^\#) \leq \liminf_{x \in B} (\langle x \mid x^\# \rangle - f(x)) + \varepsilon.$$

Citing the Cauchy principle, we are done.

Proceed with demonstrating the implication to the left. For this purpose, it suffices to take some infinitesimal $\delta > 0$ and choose $B \in \mathcal{B}$ so as to have

$$\liminf_{x \in B} (\langle x \mid x^\# \rangle - f(x)) \leq \inf(x^*(B) - f(B)) + \delta.$$

We may now refer to 5.6.18. ▷

5.6.21. Let Z be a standard Kantorovich space. Assume further that $C : E \rightarrow Z^\bullet$ is an increasing convex operator. If the sets $X \times \text{epi}(G)$ and $\text{epi}(F) \times Z$ are in general position then, given a filterbase \mathcal{B} on $\text{dom}(F)$, we have

$$D(G \circ F)(\mathcal{B}) = \bigcup_{S \in DG(F(\mathcal{B}))} D(S \circ F)(\mathcal{B}).$$

◁ The proof consists in demonstrating two inclusions. To show one of them, take $S \in DG(F(\mathcal{B}))$ and $T \in D(S \circ F)(\mathcal{B})$. By 5.6.18, we then have

$$\begin{aligned} & (\exists B' \in \mathcal{B})(\forall \bar{x} \in B')(\exists \varepsilon \in \mu(\mathcal{E})) T \in \partial^\varepsilon(S \circ F)(\bar{x}); \\ & (\exists B'' \in \mathcal{B})(\forall \bar{x} \in B'')(\exists \delta \in \mu(\mathcal{E})) S \in \partial^\delta G((S \circ F)(\bar{x})). \end{aligned}$$

Since \mathcal{B} is a filterbase, $B \subset B' \cap B''$ with some $B \in \mathcal{B}$. Moreover, for $\bar{x} \in B$, the inequalities hold:

$$(G \circ F)^*(T) \leq G^*(S) + (S \circ F)^*(T) \leq T\bar{x} - G(F(\bar{x})) + \varepsilon + \delta.$$

We here take account of a suitable rule for calculating the Young–Fenchel transform. Infinitesimals comprise a cone. Hence $\varepsilon + \delta \approx 0$ and an appeal to 5.6.18 guarantees the membership $T \in D(G \circ F)(\mathcal{B})$. Consequently, the set on the right side of the inclusion we prove lies in the set on the left side.

To show the other inclusion which is unproven yet, take $T \in D(G \circ F)(\mathcal{B})$. By 5.6.18,

$$(\forall \bar{x} \in B) (\exists \varepsilon \in \mu(\mathcal{E})) \quad (G \circ F)^*(T) \leq T\bar{x} - G(F(\bar{x})) + \varepsilon$$

for some B in \mathcal{B} . Applying an appropriate exact formula for the Young–Fenchel transform, find a positive operator $S \in L^+(Y, Z)$ satisfying

$$(G \circ F)^*(T) = G^*(S) + (S \circ F)^*(T).$$

Given $\bar{x} \in \overline{B}$, put

$$\varepsilon_1 := G^*(S) - (SF(\bar{x}) - G \circ F(\bar{x})); \quad \varepsilon_2 := (S \circ F)^*(T) - (T\bar{x} - SF\bar{x}).$$

It is clear that $0 \leq \varepsilon_1 + \varepsilon_2 \leq \varepsilon$. Therefore, ε_1 and ε_2 are infinitesimals. Finally, $S \in DG(F(B)) \subset DG(F(\mathcal{B}))$ and $T \in D(S \circ F)(B) \subset D(S \circ F)(\mathcal{B})$. \triangleright

5.6.22. Let $F_1, \dots, F_n : X \rightarrow Y^\bullet$ be convex operators, with n a standard integer. If F_1, \dots, F_n are in general position and \mathcal{B} a filterbase on $\text{dom}(F_1) \cap \dots \cap \text{dom}(F_n)$ then

$$D(F_1 + \dots + F_n)(\mathcal{B}) = D(F_1)(\mathcal{B}) + \dots + D(F_n)(\mathcal{B}).$$

\triangleleft If $T_k \in D(F_k(\mathcal{B}))$ then there are $B_1, \dots, B_n \in \mathcal{B}$ such that $T_k \in \partial^{\varepsilon_k}(F_k)(\bar{x})$ for each \bar{x} in B_k and some infinitesimal ε_k . Now, if $\bar{x} \in B_1 \cap \dots \cap B_n$ then

$$T_1 + \dots + T_n \in \partial_{\varepsilon_1 + \dots + \varepsilon_n}(F_1 + \dots + F_n)(\bar{x}).$$

The sum of standardly many infinitesimals is again an infinitesimal. An appeal to 5.6.18 corroborates the fact that the set on the right side of the equality under proof lies in the set on the left side.

Now, take $T \in D(F_1 + \dots + F_n)(\mathcal{B})$. Involving 5.6.18, observe that

$$(\forall \bar{x} \in \mathcal{B})(\exists \varepsilon \in \mu(\mathcal{E})) T \in \partial^\varepsilon(F_1 + \dots + F_n)(\bar{x})$$

for some $B \in \mathcal{B}$. Given $\bar{x} \in B$, we may thus choose an infinitesimal ε so as to have

$$(F_1 + \dots + F_n)^*(T) \leq T\bar{x} - (F_1 + \dots + F_n)(\bar{x}) + \varepsilon.$$

Using an exact formula for the Young–Fenchel transform, find T_1, \dots, T_n in $L(X, Y)$ satisfying

$$T = \sum_{k=1}^n T_k; \quad \left(\sum_{k=1}^n F_k \right)^* (T) = \sum_{k=1}^n F_k^*(T_k).$$

We now put

$$\varepsilon_k := F_k^*(T_k) - (T_k \bar{x} - F_k \bar{x}) \quad (k := 1, \dots, n).$$

It is clear that $\varepsilon_k \geq 0$ and $\varepsilon_1 + \dots + \varepsilon_n \leq \varepsilon$. Consequently, $\varepsilon_k \approx 0$ and $T_k \in DF_k(\bar{x})$ for all $k := 1, \dots, n$. This completes the proof. \triangleright

5.6.23. Let $F_1, \dots, F_n : X \rightarrow Y^\bullet$ be convex operators, with n a standard integer. Assume further that F_1, \dots, F_n are in general position, Y is a vector lattice, and \mathcal{B} a filterbase on $\text{dom}(F_1 \vee \dots \vee F_n)$. If Z is a standard Kantorovich space and T , a member of $L(Y, Z)$, is a positive linear operator then $S \in L(X, Z)$ is an infinitesimal subgradient of $T \circ (F_1 \vee \dots \vee F_n)$ along \mathcal{B} if and only if, for some $B \in \mathcal{B}$, the following system of conditions is compatible:

$$\begin{aligned} T &= \sum_{k=1}^n T_k; \quad T \in L^+(Y, Z), \quad k := 1, \dots, n; \\ \sum_{k=1}^n T_k(F_k(\bar{x})) &\approx T(F_1(\bar{x}) \vee \dots \vee F_n(\bar{x})) \quad (\bar{x} \in B); \\ S &\in \sum_{k=1}^n D(T_k \circ F_k)(B). \end{aligned}$$

\triangleleft Define the operators

$$\begin{aligned} (F_1, \dots, F_n) : X &\rightarrow (Y^n)^\bullet; \quad (F_1, \dots, F_n)(x) := (F_1(x), \dots, F_n(x)); \\ \varkappa : Y^n &\rightarrow Y; \quad \varkappa(y_1, \dots, y_n) := y_1 \vee \dots \vee y_n. \end{aligned}$$

We then have

$$T \circ F_1 \vee \dots \vee F_n = T \circ \varkappa \circ (F_1, \dots, F_n).$$

Considering 5.6.22, we finish the proof. \triangleright

5.6.24. Let X be a vector space. Assume further that Y is some Kantorovich space and \mathcal{A} is a weakly order bounded subset of $L(X, Y)$. Consider a regular convex operator $F = \varepsilon_{\mathcal{A}} \circ \langle \mathcal{A} \rangle_y$, where $\varepsilon_{\mathcal{A}}$ is as usual the canonical sublinear operator

$$\varepsilon_{\mathcal{A}} : l_\infty(\mathcal{A}, E) \rightarrow E, \quad \varepsilon_{\mathcal{A}} f := \sup f(\mathcal{A}) \quad (f \in l_\infty(\mathcal{A}, E))$$

and the affine operator $\langle \mathcal{A} \rangle_y$ with $y \in l_\infty(\mathcal{A}, y)$ acts by the rule

$$\langle \mathcal{A} \rangle_y x := \langle \mathcal{A} \rangle x + y; \quad \langle A \rangle x := A \in \mathcal{A} \mapsto Ax.$$

◁ Subdifferential calculus implies that the compatibility of the above system amounts to the membership $S \in D(G \circ F)(\bar{x})$ for all $\bar{x} \in B$. We are thus left with demonstrating the reverse implication.

Clearly,

$$D(G \circ F)(\mathcal{B}) = D(G \circ \mathcal{E}_{\mathcal{A}} \circ \langle \mathcal{A} \rangle_y) = \partial(G \circ \mathcal{E}_{\mathcal{A}})(\overline{\mathcal{B}}) \circ \langle \mathcal{A} \rangle,$$

where $\overline{\mathcal{B}} := \langle \mathcal{A} \rangle_y(\mathcal{B})$. It suffices so to represent $T \in D(G \circ \mathcal{E}_{\mathcal{A}})(\overline{\mathcal{B}})$. To this end, assume that

$$\begin{aligned} & (\exists B \in \mathcal{B}) (\forall \bar{x} \in B) (\exists \varepsilon \in \mu(\mathcal{E})) \\ & (G \circ \mathcal{E}_{\mathcal{A}})^* T \leq T \circ \langle \mathcal{A} \rangle_y \bar{x} - G \circ \mathcal{E}_{\mathcal{A}} \circ \langle \mathcal{A} \rangle_y \bar{x} + \varepsilon. \end{aligned}$$

By the general change-of-variable formulas for the Young–Fenchel transform (Theorem 3.7.10 in [279]), we have

$$(G \circ \mathcal{E}_{\mathcal{A}})^* T = G^*(T \circ \Delta_{\mathcal{A}}).$$

Given $\bar{x} \in B$, put $\bar{y} := \langle \mathcal{A} \rangle_y \bar{x}$ and obtain

$$\begin{aligned} \varepsilon & \geq (G \circ \mathcal{E}_{\mathcal{A}})^* T + T \circ \mathcal{E}_{\mathcal{A}} \bar{y} - T \bar{y} \\ & = G^*(T \circ \Delta_{\mathcal{A}}) + G \circ \mathcal{E}_{\mathcal{A}} \bar{y} - T \bar{y} \\ & = \sup_{y \in \text{dom}(G)} (T \circ \Delta_{\mathcal{A}} y - T \circ \Delta_{\mathcal{A}} \circ \mathcal{E}_{\mathcal{A}} \circ \mathcal{E}_{\mathcal{A}} \bar{y}) \\ & \quad - (G y - G \circ \mathcal{E}_{\mathcal{A}} \bar{y}) + T \circ \Delta_{\mathcal{A}} \circ \mathcal{E}_{\mathcal{A}} \bar{y} - T \bar{y} \geq 0. \end{aligned}$$

Hence, $T \circ \Delta_{\mathcal{A}} \in DG(F(B))$ and $T \circ \Delta_{\mathcal{A}} F \bar{x} \approx T \circ \langle \mathcal{A} \rangle_y \bar{x}$. The proof is complete. ▷

5.6.25. The crux of the above scheme is as follows: We find the sought presentation of a subgradient independently, in a sense, of the choice of a point, using exactness of the rules for calculating the Young–Fenchel transform. In other words, the behavior of infinitesimal subgradients analogous to conventional “exact” subgradients strongly resembles that of the ε -subdifferentials reflecting the structure of the operator under study “in the large,” i.e. globally, on the whole domain of definition. In consequence, although the rules for calculating infinitesimal subdifferentials mimic the routine formulas of local subdifferentiation, the conditions for them to hold are noticeably more stringent, coinciding with those “in the large” for the Young–Fenchel transform or for ε -subdifferentials.

Our scheme also opens an opportunity to abstract the entire spectrum of the rules for subdifferential calculus (disintegration, Rockafellar’s convolution, level sets, etc.).

5.7. Infinitesimal Optimality

In this section we study a new concept of solution to an extremal problem which rests on infinitesimals. To simplify exposition, we address only the case of “pointwise.”

5.7.1. A point $\bar{x} \in \text{dom}(f)$ is an *infinitesimal solution* of an unconstrained program $f(x) \rightarrow \inf$, with $f : X \rightarrow E^\bullet$, provided that $0 \in Df(\bar{x})$, i.e. \bar{x} is a feasible solution and $f(\bar{x}) \approx \inf\{f(x) : x \in X\}$. An infinitesimal solution to an arbitrary program is defined by analogy.

5.7.2. A standard unconstrained program $f(x) \rightarrow \inf$ has an infinitesimal solution if and only if, first, $f(X)$ is bounded below and, second, there exists a standard generalized solution $(x_\varepsilon)_{\varepsilon \in \mathcal{E}}$ of the program under consideration, i.e., $x_\varepsilon \in \text{dom}(f)$ and $f(x_\varepsilon) \leq e + \varepsilon$ for all $\varepsilon \in \mathcal{E}$, where $e := \inf f(X)$ is the value of the program.

◁ By transfer, idealization, and 5.6.3, infer that

$$\begin{aligned} (\exists \bar{x} \in X) 0 \in Df(\bar{x}) &\leftrightarrow (\exists x \in X)(\forall^{\text{st}} \varepsilon \in \mathcal{E})(0 \in \partial_\varepsilon f(x)) \\ &\leftrightarrow (\forall^{\text{st fin}} \mathcal{E}_0 \subset \mathcal{E})(\exists x \in X)(\forall \varepsilon \in \mathcal{E}_0)(0 \in \partial_\varepsilon f(x)) \\ &\leftrightarrow (\forall^{\text{st}} \varepsilon \in \mathcal{E})(\exists x_\varepsilon \in X)(0 \in \partial_\varepsilon f(x_\varepsilon)) \\ &\leftrightarrow (\forall \varepsilon \in \mathcal{E})(\exists x_\varepsilon \in X)(\forall x \in X)(f(x) \geq f(x_\varepsilon) - \varepsilon), \end{aligned}$$

which completes the proof. ▷

5.7.3. Consider a regular convex program

$$g(x) \leq 0, \quad f(x) \rightarrow \inf.$$

Thus, $g, f : X \rightarrow E^\bullet$ (for simplicity $\text{dom}(f) = \text{dom}(g) = X$), for every $x \in X$ either $g(x) \leq 0$ or $g(x) \geq 0$, and the element $g(\bar{x})$ with some $\bar{x} \in X$ is an order unit in E .

5.7.4. Assuming the standard environment, a feasible internal point x_0 is an infinitesimal solution to the regular program 5.7.3 if and only if the following system of conditions is compatible:

$$\begin{aligned} \alpha, \beta \in {}^\circ[0, \mathbf{1}_E], \quad \alpha + \beta = \mathbf{1}_E, \quad \ker(\alpha) = 0; \\ \beta \circ g(\bar{x}) \approx 0, \quad 0 \in D(\alpha \circ f)(\bar{x}) + D(\beta \circ g)(\bar{x}). \end{aligned}$$

◁ ←: In case of compatibility of the system, for a feasible x and some infinitesimals ε_1 and ε_2 we have

$$\alpha f(\bar{x}) \leq \alpha f(x) + \beta g(x) - \beta g(\bar{x}) + \varepsilon_1 + \varepsilon_2 \leq \alpha f(x) + \varepsilon$$

for every standard $\varepsilon \in {}^\circ\mathcal{E}$. In particular, $\alpha(f(\bar{x}) - f(x)) \leq \alpha\varepsilon$ for $\varepsilon \in {}^\circ\mathcal{E}$, since α is a standard mapping. Using the condition $\ker(\alpha) = 0$ and general properties of multipliers, we see that \bar{x} is an infinitesimal solution.

→: Let

$$e := \inf \{f(x) : x \in X, g(x) \leq 0\}$$

be the value of the program under consideration. By hypothesis and transfer, e is a standard element. Using the transfer principle again, by the vector minimax theorem [279, Theorem 4.1.10 (2)], we find standard multipliers $\alpha, \beta \in {}^\circ[0, \mathbb{1}_E]$ such that

$$\alpha + \beta = \mathbb{1}_E,$$

$$0 = \inf_{x \in X} \{\alpha(f(x) - e) + \beta \circ g(x)\}.$$

Arguing in a standard way, we check that $\ker(\alpha) = 0$. Moreover, since \bar{x} is an infinitesimal solution; therefore, $f(\bar{x}) - e = \varepsilon$ for some infinitesimal ε . Consequently,

$$-\alpha\varepsilon \leq \alpha f(x) - \alpha f(\bar{x}) + \beta g(x)$$

for all $x \in X$. In particular, $0 \geq \beta g(\bar{x}) \geq -\alpha\varepsilon \geq -\varepsilon$, i.e., $\beta g(\bar{x}) \approx 0$ and

$$0 \in \partial_{\alpha\varepsilon + \beta g(\bar{x})}(\alpha \circ f + \beta \circ g)(\bar{x}) \subset D(\alpha \circ f + \beta \circ g)(\bar{x}),$$

since $\alpha\varepsilon + \beta g(\bar{x}) \approx 0$. \triangleright

5.7.5. Consider a Slater regular program

$$Ax = A\bar{x}, \quad g(x) \leq 0, \quad f(x) \rightarrow \inf;$$

i.e., first, $A \in L(X, \mathcal{X})$ is a linear operator with values in some vector space \mathcal{X} , the mappings $f : X \rightarrow E^\bullet$ and $g : X \rightarrow F^\bullet$ are convex operators (for the sake of convenience we assume $\text{dom}(f) = \text{dom}(g) = X$); second, F is an Archimedean ordered vector space, E is a standard Kantorovich space of bounded elements; and, at last, the element $g(\bar{x})$ with some feasible point \bar{x} is a strong order unit in F .

5.7.6. Infinitesimal Optimality Criterion. *A feasible point \bar{x} is an infinitesimal solution of a Slater regular program if and only if the following system of conditions is compatible*

$$\begin{aligned} \gamma \in L^+(F, E), \quad \mu \in L(\mathcal{X}, E), \quad \gamma g(\bar{x}) \approx 0, \\ 0 \in Df(\bar{x}) + D(\gamma \circ g)(\bar{x}) + \mu \circ A. \end{aligned}$$

◁ ←: In case of compatibility of the system, for every feasible point x and some infinitesimals ε_1 and ε_2 , we have

$$\begin{aligned} f(\bar{x}) &\leq f(x) + \varepsilon_1 + \gamma g(x) - \gamma g(\bar{x}) + \varepsilon_2 - \mu(Ax) + \mu(A\bar{x}) \\ &\leq f(x) + \varepsilon_1 + \varepsilon_2 - \gamma g(\bar{x}) \leq f(x) + \varepsilon \end{aligned}$$

for every standard $\varepsilon \in {}^\circ\mathcal{E}$.

→: If \bar{x} is an infinitesimal solution, then \bar{x} is also an ε -solution for an appropriate infinitesimal ε . It remains to appeal to the corresponding ε -optimality criterion. ▷

5.7.7. A feasible point \bar{x} is an *infinitesimal Pareto solution* to 5.7.5 provided that \bar{x} is a Pareto ε -solution for some infinitesimal ε (with respect to the strong order unit $\mathbb{1}_E$ of the space E), i.e., if $f(x) - f(\bar{x}) \leq -\varepsilon\mathbb{1}_E$ for a feasible x then $f(x) - f(\bar{x}) = \varepsilon\mathbb{1}_E$ for $\varepsilon \in \mu(\mathbb{R}_+)$.

5.7.8. Suppose that \bar{x} is an infinitesimal Pareto solution to a Slater regular program. Then the following system of conditions is compatible for some linear functionals α , β , and γ on the spaces E , F , and \mathcal{X} respectively:

$$\begin{aligned} \alpha > 0, \quad \beta \geq 0, \quad \beta g(\bar{x}) \approx 0, \\ 0 \in D(\alpha \circ f)(\bar{x}) + D(\beta \circ g)(\bar{x}) + \gamma \circ A. \end{aligned}$$

If, in turn, the above relations are valid for some feasible point \bar{x} , $\alpha(\mathbb{1}_E) = \mathbb{1}$ and $\ker(\alpha) \cap E^+ = 0$, then \bar{x} is an infinitesimal Pareto solution.

◁ The first part of the claim ensues from the usual Pareto ε -optimality criterion on considering the basic properties of infinitesimals.

Now, if the hypothesis of the second part of the claim is valid then, appealing to the definitions, for every feasible $x \in X$ we derive

$$\begin{aligned} 0 &\leq \alpha(f(x) - f(\bar{x})) + \beta g(x) - \beta g(\bar{x}) + \varepsilon_1 + \varepsilon_2 \\ &\leq \alpha(f(x) - f(\bar{x})) + \varepsilon_1 + \varepsilon_2 - \beta g(\bar{x}) \end{aligned}$$

for appropriate infinitesimals ε_1 and ε_2 . Put $\varepsilon := \varepsilon_1 + \varepsilon_2 - \beta g(\bar{x})$. It is clear that $\varepsilon \approx 0$ and $\varepsilon \geq 0$. Now if $f(x) - f(\bar{x}) \leq -\varepsilon\mathbb{1}_E$ for a feasible x , then we obtain the equality $\alpha(f(\bar{x}) - f(x)) = \varepsilon$. In other words, $\alpha(f(\bar{x}) - f(x) - \varepsilon\mathbb{1}_E) = 0$ and $f(\bar{x}) - f(x) = \varepsilon\mathbb{1}_E$. This means that \bar{x} is a Pareto ε -optimal solution. ▷

5.7.9. Following the above pattern, we may find some tests for infinitesimal solutions to other basic types of convex programs. We leave these as exercise for the curious reader.

5.7.10. COMMENTS.

(1) A complete guide to the problems of nonsmooth analysis we deal with in this chapter is impossible to submit since this theme is very topical and even grandiose. We thus supply the reader with a few standard references, while indicating [279] as a source of details. The following brilliant articles have determined many principal directions of research in this huge area: [23, 24, 61, 80, 325, 423, 424].

(2) The Renaissance of the theory of local approximation is connected with Clarke's discovery of the convex tangent cone which is now named after him (see [60, 61]).

Clarke analyzed only the finite-dimensional case. The invention of a general definition in an abstract topological vector space turned out a delicate problem which was solved by Rockafellar. The radical changes in nonsmooth analysis that were initiated by the Clarke cone are mirrored in dozens of surveys and monographs. We list a few of them [23, 61, 79, 80, 325].

(3) Diversity of tangent cones raised the problem of their classification. Among the pilot studies in this area we must mention the articles of Dolecki [91, 92] and Ward [517–519]. The classification of tangents of this chapter belongs to Kutateladze [283, 285, 291].

(4) The regularizing cones of the types R^1 and Q^1 were suggested by Kusraev [256, 257, 260] and Thibault [485, 486].

(5) Epiconvergence theory for correspondences is due to optimization. The book [22] by Attouch played a significant role in propounding this theory. Our exposition mostly follows [291].

(6) The idea to choose special collections of infinitesimals for constructing tangents was suggested in [290]. Discussing the problems of the Cornet Theorem we mainly proceed in the wake of Hiriart-Urruty [177].

A general approach to approximation of sums and composites was suggested in [260, 262]. Our exposition follows Kutateladze [291].

(7) Infinitesimal subdifferentials appeared in a series of articles by Kutateladze. We mention only the first complete exposition of the basics of this theory in [286].

Chapter 6

Technique of Hyperapproximation

The most important applications of infinitesimal analysis to studying continuous and other infinite objects result from “discretizing” them. This implies a search for finite or at least finite-dimensional approximants that are infinitely close in some sense to the original objects. The analogy with the ubiquitous sequential schemes suggests that the “finiteness” in these approaches should involve actual infinities.

Among new entities we must first of all mention the nonstandard hulls of normed spaces which were discovered by Luxemburg and the class of measures which were introduced by Loeb and are known now as Loeb measures in common parlance. The concept of nonstandard hull enables us to model infinite-dimensional Banach spaces with hyperfinite-dimensional spaces. The construction of Loeb measure enables us to model the conventional measure spaces with measures on hyperfinite sets. We combine these ideas in the term “hyperapproximation.”

In this chapter we consider these constructions and their applications to discrete approximation of Banach spaces and hyperapproximation of integral and pseudo-integral operators.

Throughout the chapter we work in the framework of the classical and even radical stances of infinitesimal analysis since the constructions we use often manipulate all types of cantor sets available in infinitesimal analysis. This in particular involves the necessity of cluttering the language of exposition with the prefix “hyper” for distinguishing between standard and nonstandard finite sets, standard and nonstandard finite-dimensional vector spaces, etc.

The reader should bear in mind that some type of bilingualism is absolutely unavoidable in confessing whatever stance of infinitesimal analysis.

Without further specification, we always choose some nonstandard universe that is sufficiently representative for our needs or work in an appropriate fragment of such a universe (a superstructure in nonstandard parlance). We always presume that the necessary form of saturation or \aleph -saturation is satisfied in the ambient universe of discourse. We freely use the strong concurrence principle or Nelson’s

idealization, answering the challenges of the context (cf. 3.5.2–3.5.11). It is worth observing that this freedom of action is absolutely lawful, revealing the inestimable advantages of the “nonstandard” practice of doing mathematics.

6.1. Nonstandard Hulls

In this section we describe an important construction of infinitesimal analysis which serves as one of the most useful tools of further exposition.

6.1.1. Suppose that E is an internal vector space over ${}^*\mathbb{F}$ where \mathbb{F} is a *basic field* of scalars, i.e. \mathbb{R} or \mathbb{C} . In other words, given are the two internal operations $+ : E \times E \rightarrow E$ and $\cdot : {}^*\mathbb{F} \times E \rightarrow E$ satisfying the usual axioms of vector spaces. Since $\mathbb{F} \subset {}^*\mathbb{F}$; therefore, each internal vector space E is simultaneously an external vector space over \mathbb{F} but neither a normed nor a Hilbert space even if E is such in the internal universe. Nevertheless, it is possible to associate with each internal normed or pre-Hilbert some external Banach or Hilbert space.

Let $(E, \|\cdot\|)$ be an internal normed space over ${}^*\mathbb{F}$. As usual, an element x in E is *limited* or *infinitesimal* provided that so is $\|x\|$. Denote by $\text{ltd}(E)$ and $\mu(E)$ the external sets of limited and infinitesimal elements of a normed space E . The notation $\mu(E)$ conforms with the agreements of Chapter 4 because $\mu(E)$ coincides with the monad of the neighborhood filter of the zero of E .

Clearly, $\text{ltd}(E)$ is an external vector space over \mathbb{F} , and $\mu(E)$ is a subspace of $\text{ltd}(E)$. Denote the quotient space $\text{ltd}(E)/\mu(E)$ by $E^\#$ and make $E^\#$ a normed space by putting

$$\|\pi x\| := \|x^\#\| := \text{st}(\|x\|) \in \mathbb{F} \quad (x \in \text{ltd}(E)),$$

where $\pi := (\cdot)^\# : \text{ltd}(E) \rightarrow E^\#$ is the quotient mapping. The external normed space $(E^\#, \|\cdot\|)$ is the *nonstandard hull* of E . If $(E, \|\cdot\|)$ is a standard normed space then the nonstandard hull of E is by definition the space $({}^*E)^\#$.

If $x \in E$ then $\pi({}^*x) = ({}^*x)^\#$ is a member of $({}^*E)^\#$ and, moreover, $\|x\| = \|({}^*x)^\#\|$. Thus, $x \mapsto ({}^*x)^\#$ is an isometric embedding of E into $({}^*E)^\#$. The image of E under this embedding is always identified with E , providing the inclusion $E \subset ({}^*E)^\#$.

Suppose now that E and F are internal normed spaces and $T : E \rightarrow F$ is an internal bounded linear operator. Note that

$$c(T) := \{C \in {}^*\mathbb{R} : (\forall x \in E) \|Tx\| \leq C\|x\|\}$$

is an internal bounded set of reals. Recall that $\|T\| := \inf c(T)$ by the definition of *operator norm*.

If $\|T\|$ is a limited real, then it follows from the classical normative inequality $\|Tx\| \leq \|T\| \|x\|$, valid for all $x \in E$, that $T(\text{ltd}(E)) \subset \text{ltd}(F)$ and $T(\mu(E)) \subset \mu(F)$. Therefore, the rule

$$T^\# \pi x := \pi T x \quad (x \in E)$$

soundly defines the external operator $T^\# : E^\# \rightarrow F^\#$ that is the quotient of T by the infinite proximity on E . The operator $T^\#$ is \mathbb{F} -linear and bounded; moreover, $\|T^\#\| = \text{st}(\|T\|)$. It is in common parlance to call $T^\#$ the *nonstandard hull* of T .

6.1.2. Theorem. *The nonstandard hull $E^\#$ is a Banach space for each internal (not necessarily complete) normed space E .*

◁ Recall the definitions and notation for the open and closed balls with center a and radius ε : If X is a normed space, $x \in X$, and $\varepsilon > 0$ then

$$\begin{aligned} \overset{\circ}{B}_X(a, \varepsilon) &:= \overset{\circ}{B}_\varepsilon(a) := \{x \in X : \|x - a\| < 1\}, \\ B_X(a, \varepsilon) &:= B_\varepsilon(a) := \{x \in X : \|x - a\| \leq 1\}. \end{aligned}$$

As suggested by the nested ball test for completeness (see, for instance, [300, 4.5.7]), take some sequence of balls $B_{E^\#}(x_n^\#, r_n)$, with $x_n \in E$, $r_n \in \mathbb{R}$ for all $n \in \mathbb{N}$, and $\lim_{n \rightarrow \infty} r_n = 0$. We may assume that (r_n) decreases. Consider the nested sequence of internal closed balls $B_E(x_n, r_n + r_n/2^{n+1})$ in E . By saturation, there is some $x \in E$ belonging to each of these balls. The element $x^\#$ is a common point of the balls $B_{E^\#}(x_n^\#, r_n)$ for $n \in \mathbb{N}$. ▷

6.1.3. Let E be an internal vector space whose dimension is limited. In this event E is referred to as *hyperfinite-dimensional*. Internal hyperfinite-dimensional spaces will play a key role in what follows and so it is worthwhile to elaborate a few relevant details.

We first clarify the concept of the sum of hyperfinitely many elements of a vector space. To this end, rephrase the definition of the sum of finitely many vectors as a set-theoretic formula and proceed by transfer.

In more detail, if f is a sequence in E (i.e., $f : \mathbb{N} \rightarrow E$) then the partial sums $g(n)$ of the series $\sum_{k=0}^\infty f(k)$ are defined by recursion as follows:

$$\text{Seq}(f) \wedge \text{Seq}(g) \wedge f(0) = g(0) \wedge (\forall k \in \mathbb{N})(g(k+1) = g(k) + f(k+1)),$$

with $\text{Seq}(g)$ standing for the formal record of the formula: “ g is a sequence.”

Abbreviate the above formula to $\Sigma(f, g)$ and denote the cardinality of a set M by $|M|$. We also identify a natural k with the set $\{0, 1, \dots, k-1\}$. Take as E an internal vector space (or an internal abelian group) and let Y be a hyperfinite

subset of E . Distinguish some bijection $f : \{0, \dots, |Y| - 1\} \rightarrow Y$ and extend f to the internal sequence $f : {}^*\mathbb{N} \rightarrow Y$ by letting f be zero for $n > |Y| - 1$. Proceed with defining the sequence $g : {}^*\mathbb{N} \rightarrow E$ by using $\Sigma(f, g)$.

By transfer, $g(|Y| - 1)$ is independent of the choice of f , and so we may soundly put $\sum_{x \in Y} x := g(|Y| - 1)$. Appealing to the transfer principle again, note that the so-defined sum of a hyperfinite set enjoys all properties of the sum of finitely many addends. For example, if $\{Y_m : m < \nu\}$ is a hyperfinite internal family of subsets of Y presenting a partition of Y then

$$\sum_{x \in Y} x = \sum_{k=0}^{\nu-1} \sum_{x \in Y_k} x.$$

It is now easy to define an internal vector space of hyperfinite dimension as follows:

Let E be an internal linear space. A hyperfinite internal set $\{e_1, \dots, e_\Omega\}$, with $\Omega \in {}^*\mathbb{N}$, is a *basis* for E provided that to each $x \in E$ there is a unique hyperfinite internal set $\{x_1, \dots, x_\Omega\}$ in ${}^*\mathbb{F}$ satisfying $x = \sum_{k=1}^{\Omega} x_k e_k$. If E has a hyperfinite basis then E is called *hyperfinite-dimensional*, and the internal cardinality Ω of this basis is the *internal dimension* of E ; in symbols, $\dim(E) := \Omega$.

By transfer, all properties of finite-dimensional vector spaces and their finite bases carry over to hyperfinite-dimensional vector spaces and their hyperfinite bases. For example, $\dim(E) = \Omega$ if and only if there exists an internally linearly independent hyperfinite subset Y in E of internal cardinality Ω , and every hyperfinite set of internal cardinality $\Omega + 1$ is internally linearly dependent.

Clearly, we say that some hyperfinite internal set $\{y_1, \dots, y_\nu\}$ is *internally linearly independent* provided that $\sum_{j=0}^{\nu} \lambda_j y_j \neq 0$ for every internal finite sequence $\{\lambda_1, \dots, \lambda_\nu\}$ with at least one nonzero element.

Note that if a set $\{y_1, \dots, y_\nu\}$ is internally linearly independent, then it is linearly independent also in the external universe. Indeed, the external linear dependence property of E means the linear dependence property over \mathbb{F} for all standardly finite subsets of E , and the latter is still linear dependence in the internal universe since every standardly finite set is internal and $\mathbb{F} \subset {}^*\mathbb{F}$.

On the other hand, a linearly independent set in the internal universe may fail to be so externally. For instance, if $x \in E$, $x \neq 0$, and $\alpha \in {}^*\mathbb{F} - \mathbb{F}$ then $\{x, \alpha x\}$ is a linearly independent set internally but it is linearly dependent in the external universe since $\alpha \notin \mathbb{F}$.

Speaking about internal vector spaces, we will imply the interval versions of linear dependence and independence, dimension, and so on. Therefore, the adjective “internal” itself will be omitted as a rule.

6.1.4. The most typical example of a hyperfinite-dimensional vector space is the space consisting of all internal mappings $x : T \rightarrow {}^*\mathbb{C}$ from some hyperfinite set T to the field of hypercomplex numbers ${}^*\mathbb{C}$. The resulting vector space is furnished with the internal norm

$$\|x\|_p := \left(\sum_{t \in T} |x(t)|^p \right)^{1/p} \quad (x \in ({}^*\mathbb{C})^T),$$

with $p \in {}^*\mathbb{R}$, $1 \leq p$, and denoted by $l_p(T)$ or $l_p(n)$, with n standing for the size of T .

It is in common parlance to consider the internal inner product

$$\langle x, y \rangle := \sum_{t \in T} x(t) \overline{y(t)}$$

giving rise to Hilbert norm of x coincident with $\|x\|_2$; moreover, the index $p = 2$ is omitted in the notation for the norm.

In the internal universe all norms on $({}^*\mathbb{C})^T$ are equivalent by transfer, i.e., $(\exists C_1, C_2 > 0)(\forall x)(C_1\|x\|_{p_1} \leq \|x\|_{p_2} \leq C_2\|x\|_{p_1})$. It is worth observing that the constants C_1 and C_2 belong to ${}^*\mathbb{R}$ and so they might be infinitesimal or unlimited sometimes.

Recall that if E is a normed space (over \mathbb{C} for definiteness) then E is usually treated as a topological space in which the neighborhood filter of an arbitrary point x has the shape

$$\tau_E(x) := \text{fil}\{B_\varepsilon(x) : \varepsilon \in \mathbb{R}_+\}.$$

In this event, *E is an internal normed space over ${}^*\mathbb{C}$ and the neighborhood filter of an arbitrary point y of *E has the shape

$$\tau_{{}^*E}(y) := \text{fil}\{B_\varepsilon(y) : \varepsilon \in {}^*\mathbb{R}_+\}.$$

We usually drop $*$ while indicating the addition and scalar multiplication as well as norm and inner product of *E .

The space $l_p(n)$ is an internal Banach lattice. We note for the sake of completeness that if E is an internal normed vector lattice then $E^\#$ is naturally equipped with some order induced by the quotient mapping $x \mapsto x^\#$. Namely, the positive cone of $E^\#$ is defined as $E^\# := \{x^\# : 0 \leq x \in \text{ld}(E)\}$. Moreover:

The nonstandard hull $E^\#$ of an internal normed vector lattice E is a Banach lattice with sequentially order continuous norm. Moreover, every increasing and norm bounded sequence in $E^\#$ is order bounded.

If p and n are limited numbers then $l_p(n)^\#$ is a Banach lattice order isomorphic and isometric to $l_q(n)$, where $q := \text{st}(p)$ (see 6.1.5). Assume further that p is unlimited whereas n remains limited. Then the space $l_p(n)^\#$ is isomorphic to $l_\infty(n)$. It is possible to show also that, in the case of n unlimited and $p \geq 1$ limited, $l_p(n)$ is isomorphic to $L_q(\mu)$ for some measure μ . In case both numbers n and $p \geq 1$ are unlimited, $l_p(n)$ is isomorphic to $L_\infty(\mu)$.

6.1.5. Theorem. *If E is an internal finite-dimensional normed space and $n := \dim(E)$ is standard then $\dim(E^\#) = n$. Otherwise, $E^\#$ is nonseparable.*

◁ Check first that if $E^\#$ is separable then $E^\#$ is finite-dimensional. To this end, put $A^\# := \{e^\# : e \in A\} \subset E^\#$ for $A \subset E$. This notation applies both to external and internal subsets of E . If X is a normed space then we let B_X stand for the *unit ball* of X , i.e. $B_X := \{x \in X : \|x\| \leq 1\}$. Clearly, $B_{E^\#} = (B_E)^\#$.

Indeed, the inclusion $(B_E)^\# \subset B_{E^\#}$ follows from the fact that $(\forall \lambda \in {}^*\mathbb{R}) (\lambda \leq 1 \rightarrow \circ\lambda \leq 1)$. Take $\xi \in B_{E^\#}$, i.e., $\|\xi\| \leq 1$. If $\|\xi\| < 1$ then $(\exists e \in E)(\xi = e^\# \wedge \|e\| < 1)$, i.e., $e \in B_E$, and so $\xi \in (B_E)^\#$. If $\|\xi\| = 1$ then $\xi = e^\#$, where $\|e\| \approx 1$, and it may happen so that $\|e\| > 1$. However, in this case $e' := \frac{e}{\|e\|} \approx e$, implying that $e'^\# = e^\#$. Consequently, $\|e'\| = 1$, and again $\xi \in (B_E)^\#$. Similarly, if $e \in \text{ltd}(E)$ then $\overset{\circ}{B}_E(e, \varepsilon)^\# \subset B_{E^\#}(e^\#, \varepsilon)$.

To prove that $E^\#$ is finite-dimensional it suffices to show that the ball $B_{E^\#}$ is compact. Since $E^\#$ is separable by hypothesis, there is a countable dense subset $\{e_k^\# : k \in \mathbb{N}\}$ of $B_{E^\#}$. By above, we may assume that $e_k \in B_E$ for all $k \in \mathbb{N}$.

Take an arbitrary $\varepsilon > 0$ and consider the increasing sequence of internal sets $M_n := \bigcup_{k=0}^n \overset{\circ}{B}_E(e_k, \varepsilon) \cap B_E$. Show that $\bigcup_{n \in \mathbb{N}} M_n = B_E$. Indeed, if $e \in B_E$ then $\|e^\# - e_n^\#\| < \varepsilon$ for some $n \in \mathbb{N}$, implying that $\|e - e_n\| < \varepsilon$, i.e., $e \in B_E(e_n, \varepsilon) \subset M_n$. By saturation, $M_{n_0} = B_E$ for some $n_0 \in {}^*\mathbb{N}$. Therefore,

$$\begin{aligned} B_{E^\#} &= \left(\bigcup_{k=0}^{n_0} \overset{\circ}{B}_E(e_k, \varepsilon) \cap B_E \right)^\# \\ &= \bigcup_{k=0}^{n_0} \left(\overset{\circ}{B}_E(e_k, \varepsilon) \cap B_E \right)^\# \\ &\subset \bigcup_{k=0}^{n_0} \overset{\circ}{B}_{E^\#}(e_k^\#, \varepsilon) \cap B_{E^\#}. \end{aligned}$$

This inclusion allows us to conclude that $\{e_0^\#, \dots, e_{n_0}^\#\}$ is an ε -net for $B_{E^\#}$.

Show now that if $e_1, \dots, e_n \in \text{ltd}(E)$, where $n \in \mathbb{N}$ and $e_1^\#, \dots, e_n^\#$ are linearly independent, then e_1, \dots, e_n are linearly independent in E (over ${}^*\mathbb{F}$). Indeed, suppose that $\lambda_1, \dots, \lambda_n \in {}^*\mathbb{F}$ are not all zero and $\sum_{k=1}^n \lambda_k e_k = 0$. If $\lambda := \max_k |\lambda_k|$

then $\lambda \neq 0$ and $\sum_{k=1}^n \mu_k e_k = 0$, where $\mu_k := \lambda_k/\lambda$ and $|\mu_k| \leq 1$. Since $|\mu_k|$ is limited; therefore, ${}^\circ\mu_k$ lies in ${}^*\mathbb{R}$. It is easy in this case that $(\sum_{k=1}^n \mu_k e_k)^\# = \sum_{k=1}^n {}^\circ\mu_k e_k^\# = 0$. However, $\|\mu_j\| = 1$ for some j , i.e., ${}^\circ\mu_j \neq 0$, which contradicts the linear independence of $e_1^\#, \dots, e_n^\#$.

The above arguments show, in particular, that if the internal dimension of E is a standard number n , i.e., $\dim(E) = n$; then $\dim(E^\#) \leq n$. To prove the reverse inequality we use an Auerbach basis (see [170]).

A basis $\{e_1, \dots, e_n\}$ in a normed space X is an *Auerbach basis* provided that $\|e_1\| = \dots = \|e_n\| = 1$ and

$$\left\| \sum_{k=1}^n \alpha_k e_k \right\| \geq |\alpha_j| \quad (\alpha_1, \dots, \alpha_n \in \mathbb{F})$$

for all $j := 1, \dots, n$. This amounts to the fact that the linear span of the set $(e_k)_{k \neq j}$ is *orthogonal* to the vector e_j on the understanding that the *orthogonality* of $M \subset L$ to $x \in L$ means that $\|x + \alpha y\| \geq \|x\|$ for all $\alpha \in \mathbb{F}$ and $x \in M$, cf. [128].

It is well known (see, for example, [128]) that each finite-dimensional normed space has an Auerbach basis. Consequently, such an internal basis exists also in the internal n -dimensional space E by transfer.

Suppose that $\{e_1, \dots, e_n\}$ is an Auerbach basis for E . Check that $e_1^\#, \dots, e_n^\#$ are linearly independent.

If $\sum_{k=1}^n \lambda_k e_k^\# = 0$ then $\|\sum_{k=1}^n \lambda_k e_k\| \approx 0$, but $\|\sum_{k=1}^n \lambda_k e_k\| \geq |\lambda_j|$ for all $j := 1, \dots, n$, contradicting the fact that all λ_k are standard, and at least one of them is nonzero. This proves that the internal dimension of E equals n ; in symbols, $\dim(E^\#) = n$. We have so completed the proof of the first claim of the theorem.

Suppose now that the internal dimension of E is greater than each standard n . Then to such a natural n there is an internal subspace $E_1 \subset E$ satisfying $\dim(E_1) = n$. Obviously, $E_1^\#$ is embedded in $E^\#$ isometrically, and so $E^\#$ contains some n -dimensional subspace for all $n \in \mathbb{N}$. In this event $E^\#$ is neither finite-dimensional nor separable. \triangleright

6.1.6. Let $\mathfrak{F}(E)$ stand for the set of all finite-dimensional subspaces of a normed space E . Given $F \in \mathfrak{F}(E)$, denote the dimension of F by $\dim(F)$. By transfer ${}^*\mathfrak{F}(E)$ consists of some (not necessarily all) hyperfinite-dimensional subspaces of the internal space *E . Moreover, ${}^*\dim$ is a mapping from ${}^*\mathfrak{F}(E)$ to ${}^*\mathbb{N}$ satisfying ${}^*\dim(F) = \dim(F)$ for all $F \in \mathfrak{F}(E)$.

To each vector space E there is some F in ${}^\mathfrak{F}(E)$ satisfying $E \subset F \subset {}^*E$. In other words, there is some hyperfinite-dimensional subspace $F \subset {}^*E$ that contains all standard elements of the internal space *E .*

◁ The proof is straightforward by saturation. Indeed, put $A_x := \{F \in {}^*\mathfrak{F}(E) : x \in F\}$ for all $x \in E$. The family $(A_x)_{x \in E}$ of internal sets is nested and so it possesses the finite intersection property. By saturation, there is some F in ${}^*\mathfrak{F}(E)$ such that $x \in F$ for all $x \in X$. ▷

Despite simplicity of formulation and proof, the above proposition lays grounds for numerous applications of infinitesimal analysis in Banach space theory. The scheme of actions is here as follows:

Embed E into a hyperfinite-dimensional vector space F . By transfer, we may establish various facts about the space F and operators in F by proving them for finite-dimensional subspaces of E and operators in them. Since E lies in F , it follows that we may use the standard part operation to obtain results about E and the endomorphisms of E . This scheme is not automatic and sometimes, requires a rather sophisticate technique for implementation. In particular, the object under consideration may lack the nearstandardness property which is necessary for using the standard parts and so we are to introduce nearstandardness in an ad hoc manner, cf. [342].

6.1.7. We now shortly consider what happens with the property of an operator in passing to nonstandard hulls.

Let E , F , and G be internal normed spaces over the field ${}^*\mathbb{F}$ (which is the standardization of the basic field \mathbb{F} and coincides with ${}^*\mathbb{R}$ or ${}^*\mathbb{C}$). Assume that $S, T : E \rightarrow F$ and $R : F \rightarrow G$ are limited internal operators. Then

- (1) $\|T^\#\| = {}^\circ\|T\|$;
- (2) $(S + T)^\# = S^\# + T^\#$;
- (3) $(\lambda T)^\# = ({}^\circ\lambda)T^\#$ for every $\lambda \in \text{ltd}({}^*\mathbb{F})$;
- (4) $(R \circ T)^\# = R^\# \circ T^\#$.

◁ These are all obvious. ▷

6.1.8. Suppose that E is an internal vector space with the inner product $\langle \cdot, \cdot \rangle$. As mentioned above, $E^\#$ is the nonseparable Hilbert space whose inner product (\cdot, \cdot) looks as follows

$$(x^\#, y^\#) := \langle x, y \rangle \quad (x, y \in E).$$

If T is an operator between Hilbert spaces then we let T^* stand the *adjoint* of T (also known as the *hermitian conjugate* of T). Let $\sigma(T)$ be the *spectrum* of T and $\sigma_p(T)$, the *point spectrum* of T (i.e., the set of eigenvalues of T).

Assume that E is an internal pre-Hilbert space and $T : E \rightarrow E$ is an internal linear operator with limited norm. Then

- (1) $(T^*)^\# = (T^\#)^*$;
- (2) If T is hermitian, or normal, or unitary operators then so is $T^\#$;
- (3) If E is hyperfinite-dimensional then $\sigma(T^\#) = \sigma_p(T^\#)$.

6.1.9. *If E is an internal pre-Hilbert space and F is an internal subspace of E then $(F^\perp)^\# = (F^\#)^\perp$.*

◁ Let P_F be the orthoprojection in E to F , and let $P_{F^\#}$ be the orthoprojection of $E^\#$ to $F^\#$. Show that $P_F^\# = P_{F^\#}$.

By 6.1.7 and 6.1.8, $P_F^\#$ is an orthoprojection. We are left with proving that $P_F^\# \xi = \xi$ if and only if $\xi \in F^\#$.

If $\xi = x^\#$ and $x \in F$ then $P_F^\# x^\# = (P_F x)^\# = x^\# = \xi$. Conversely, suppose that $P_F^\# \xi = \xi$. If $\xi = y^\#$ then $(P_F y)^\# = y^\#$, and so $P_F y - y \approx 0$. Putting $x := P_F y$, observe that $\xi = x^\#$. Since $x = P_F y \in F$; therefore, $\xi \in F^\#$.

To complete the proof, note that $H = F^\perp$ if and only if $P_H + P_F = I$ and $P_H P_F = P_F P_H = 0$. Appealing again to 6.1.7, infer that $P_{H^\#} + P_{F^\#} = I^\#$ and $P_{H^\#} P_{F^\#} = P_{F^\#} P_{H^\#} = 0^\#$. Hence, $H^\# = (F^\#)^\perp$. ▷

We now provide three auxiliary facts helpful in the sequel. Note that in 6.1.10–6.1.12 E stands for a hyperfinite-dimensional Hilbert space.

6.1.10. *If T is an internal limited normal operator then $\sigma(T^\#) = \{\circ\lambda : \lambda \in \sigma(T)\}$.*

◁ Put $B := \{\circ\lambda : \lambda \in \sigma(T)\}$. Clearly, B is a closed subset of \mathbb{C} , since $\sigma(T)$ is internal set by 4.2.5. Obviously, $B \subset \sigma(T^\#)$. Suppose that $\mu \notin B$. Then there is a standard real $\delta > 0$ satisfying $|\mu - \xi| \geq \delta$ for all $\xi \in B$. Consequently, $|\mu - \xi| \geq \delta$ for all $\xi \in \sigma(T)$.

Since $\mu \notin \sigma(T)$; therefore, $(\mu - T)^{-1}$ is a bounded linear operator. It is clear that

$$\sigma((\mu - T)^{-1}) = \left\{ \frac{1}{\mu - \lambda} : \lambda \in \sigma(T) \right\}.$$

Since T is normal, the last equality implies that $\|(\mu - T)^{-1}\| \leq \delta^{-1}$, i.e., the norm of $(\mu - T)^{-1}$ is limited. By 6.1.7 (4), $(\mu - T)^{-1\#} = (\mu(I_E)^\# - T^\#)^{-1} = (\mu - T^\#)^{-1}$, since $(I_E)^\# = I_{E^\#}$ is the identity operator on $E^\#$. Hence, $\mu \notin \sigma(T^\#)$. ▷

6.1.11. *Assume that $\dim(E) = N \in {}^*\mathbb{N}$ and let $A : E \rightarrow E$ be an internal hermitian endomorphism whose matrix $(a_{kl})_{k,l=1}^N$ in some orthonormal basis enjoys the condition $\sum_{k,l=1}^N |a_{kl}|^2 < +\infty$. Then all eigenvalues of A are limited and the multiplicity of each noninfinitesimal eigenvalue is a standard natural.*

◁ This is immediate from the equality

$$\sum_{k=1}^s n_k |\lambda_k|^2 = \sum_{k,l=1}^N |a_{kl}|^2,$$

where $\lambda_1, \dots, \lambda_s$ is the complete list of distinct eigenvalues of A with respective multiplicities n_1, \dots, n_s . ▷

6.1.12. Let $A : E \rightarrow E$ be a bounded hermitian operator and $\lambda \in \mathbb{R}$, a standard number. Let \mathcal{M} stand for the set of eigenvalues of A , infinitely close to λ . Assume that \mathcal{M} is internal and standardly finite; i.e., $\mathcal{M} = \{\lambda_1, \dots, \lambda_n\}$ for some $n \in \mathbb{N}$. Assume further that the multiplicity of each $\lambda_k \in \mathcal{M}$ is standard, and denote by $\{\varphi_1, \dots, \varphi_m\}$ a complete orthonormal system of eigenvectors with eigenvalues in \mathcal{M} . In this case if $f \in E^\#$ is an eigenvector of $A^\#$ with eigenvalue λ then f is expressible as a linear combination of the elements $\varphi_1^\#, \dots, \varphi_m^\#$.

◁ Let H be the internal linear span of $\{\varphi_1, \dots, \varphi_m\}$. Since m is standard, it is easy to see that $H^\# \subset E^\#$ is the linear span of $\{\varphi_1^\#, \dots, \varphi_m^\#\}$. Denote by $E_\lambda^\#$ the eigenspace of $A^\#$ with eigenvalue λ . Obviously, $H^\# \subset E_\lambda^\#$. If $H^\# \neq E_\lambda^\#$ then $(H^\#)^\perp \cap E_\lambda^\# \neq 0$.

By 6.1.9, $H^{\#\perp} = (H^\perp)^\#$. By definition, H is an invariant subspace of A , and so H^\perp is an invariant subspace of A too. Hence, $(H^\perp)^\#$ is an invariant subspace of $A^\#$. Suppose that $0 \neq f \in (H^\perp)^\# \cap E_\lambda^\#$. Then λ is an eigenvalue of the restriction $A|_{H^\perp}$ of $A^\#$ to $(H^\perp)^\#$. Consequently, there is an eigenvalue γ of $A|_{H^\perp}$. Each eigenvector with eigenvalue γ is orthogonal to H and, hence, to all φ_i , which is a contradiction. ▷

6.1.13. COMMENTS.

(1) It is Luxemburg who discovered the nonstandard hull of a Banach space [328]. The ultraproducts of Banach spaces, introduced by Dacunha-Castelle and Krivine [72], are analogs of nonstandard hulls. Consult [160, 167, 170] about the role of these notions in Banach space theory and further references.

(2) The question of analytical description for nonstandard hulls, we mentioned in the end of 6.1.4, is studied in detail for the classical Banach spaces; see [64]. An arbitrary axiomatic external set theory enables us to obtain only results similar to 6.1.4. Nevertheless, it is possible to elaborate description for nonstandard hulls in a restricted fragment of the von Neumann universe. For instance, if we assume that the nonstandard universe is ω_0 -saturated (which is a constraint from below) and possesses the ω_0 -isomorphism property (which is a constraint from above) then the nonstandard hull of the Banach lattice $L_p([0, 1])$ is isometrically isomorphic with the l_p -sum of k copies of $L_p([0, 1]^k)$, with $k := 2^{\omega_0}$. See [165, 170] for a detailed exposition of this fact and further references.

(3) Recall that some properties of a normed space E are “local” in the sense that they are defined by the structure and location of finite-dimensional subspaces of E .

Nonstandard hulls have interesting local properties. For instance, it often happens so that if some property holds “approximately” on finite-dimensional subspaces then it holds “exactly” in the nonstandard hull of the ambient space. An example is the notion of finite representability, see [66, 170]. Dvoretzky had introduced the

notion of finite representability (the term was minted by James) into Banach space theory long before the model-theoretic technique has entered the latter.

(4) The problem of conditions for Banach spaces to have isomorphic nonstandard hulls was addressed by Henson [164]. Using a special first-order language, he considered the property of approximate equivalence of Banach spaces which amounts to isometric isomorphism of their nonstandard hulls [164]; also see [460, 461].

(5) Propositions 6.1.7 and 6.1.8 are established in [371]. Propositions 6.1.9–6.1.12 are taken from [146]. Proposition 6.1.10, valid for every normal operator in a hyperfinite-dimensional Hilbert space, fails in more general situations (some counterexamples are given in [528]).

6.2. Discrete Approximation in Banach Space

Study of linear operator equations in Banach space constantly and successfully use the method of discretization or “discrete approximation” which consists in replacing the original equation with some approximate equation in finite dimensions. This technique involves an important problem of the limit behavior of the spectra of approximants. The present section suggests an infinitesimal approach to this area of research.

6.2.1. We start with the definitions of discrete approximation of a Banach space and a linear operator.

Let X and X_n , with $n \in \mathbb{N}$, be some Banach spaces whose norms are denoted by $\|\cdot\|$ and $\|\cdot\|_n$. Assume given a dense subspace $Y \subset X$ and a sequence of surjective linear operators $(T_n) : Y \rightarrow X_n$ satisfying

$$(1) \quad \lim_{n \rightarrow \infty} \|T_n(f)\|_n = \|f\| \quad (f \in Y).$$

In this event $((X_n, T_n))_{n \in \mathbb{N}}$ is a *discrete approximant* to X . The term “strong discrete approximation” signifies the case in which $Y = X$.

A sequence $(x_n)_{n \in \mathbb{N}}$, with $x_n \in X_n$, *converges discretely* to f in Y provided that $\|T_n f - x_n\|_n \rightarrow 0$ as $n \rightarrow \infty$.

Let $((X_n, T_n))_{n \in \mathbb{N}}$ be a discrete approximant to X . Assume given a (possibly, unbounded) linear operator $A : X \rightarrow X$ and a sequence (A_n) , with A_n an endomorphism of X_n . Denote by $D\text{Ap}(A)$ the subspace of Y that comprises all $f \in Y$ such that $Af \in Y$ and

$$(2) \quad \lim_{n \rightarrow \infty} \|T_n A f - A_n T_n f\|_n = 0.$$

(In other words, $(A_n T_n f)$ converges discretely to Af .)

We call $D\text{Ap}(A)$ the *approximation domain* of A by (A_n) . If $D\text{Ap}(A)$ is dense in Y then we say that the sequence of operators (A_n) converges discretely to A . In case $((X_n, T_n))_{n \in \mathbb{N}}$ is a strong discrete approximant and $\|T_n A - A_n T_n\|_n \rightarrow 0$ as $n \rightarrow \infty$, we speak about *uniform discrete convergence*.

6.2.2. If $((X_n, T_n))_{n \in \mathbb{N}}$ is a strong discrete approximant then $(T_n)_{n \in \mathbb{N}}$ is a uniformly bounded sequence: $(\exists C > 0)(\forall n \in \mathbb{N})(\|T_n\|_n \leq C)$.

◁ This assertion is an instance of the celebrated boundedness principles of classical functional analysis (see, for instance, in [300, 7.2]). ▷

6.2.3. Let X and X_n be Hilbert spaces with inner products $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle_n$ respectively. Assume that $((X_n, T_n))_{n \in \mathbb{N}}$ is a discrete approximant to X . Distinguish some unlimited natural N and consider the internal Hilbert space X_N . In what follows we agree to let \mathcal{X} stand for the nonstandard hull of $X_N^\#$. We now define an embedding $\mathbf{t} : X \rightarrow \mathcal{X}$.

Let $Y \subset X$ be some dense subspace satisfying the definition of 6.2.1. Given $f \in Y$, put $\mathbf{t}(f) := T_N(f)^\#$. By 6.1.7 (1) and the infinitesimal limit test (cf. Theorem 2.3.1) $\|f\| \approx \|T_N(f)\|_N$. This obviously implies the equality $\|f\| = \|\mathbf{t}(f)\|$. Hence, $\mathbf{t} : Y \rightarrow \mathcal{X}$ is a linear isometry. Consequently, \mathbf{t} has a unique extension by continuity to the whole \mathcal{X} which we will also denote by \mathbf{t} .

Consider now a sequence of linear operators (A_n) , with A_n an endomorphism of X_n , which converges discretely to a bounded linear operator $A : X \rightarrow X$ (see the definition of discrete approximation).

Assume at first that this sequence is uniformly bounded; i.e., $(\exists C > 0)(\forall n \in \mathbb{N})(\|A_n\|_n \leq C)$. Then the internal linear operator A_N is also bounded and, moreover, its norm $\|A_N\|_N$ is a limited hyperreal. Thus A_N determines the bounded linear operator $A_N^\# : \mathcal{X} \rightarrow \mathcal{X}$ by the rule:

$$A_N^\#(x^\#) = A_N(x)^\# \quad (x \in \text{ltd}(X_N)).$$

In what follows we denote this operator by \mathcal{A} . Obviously, $\|\mathcal{A}\| = {}^\circ\|A_N\|_N$.

(1) If a sequence of linear operators (A_n) , with A_n an endomorphism of X_n , is uniformly bounded then (A_n) converges discretely to a bounded endomorphism A of X if and only if the diagram

$$\begin{array}{ccc} X & \xrightarrow{A} & X \\ \mathbf{t} \downarrow & & \downarrow \mathbf{t} \\ \mathcal{X} & \xrightarrow{\mathcal{A}} & \mathcal{X} \end{array}$$

commutes for all infinite N .

◁ From 6.2.1 (2) it follows that $\|A_N T_N f - T_N A f\|_N \approx 0$ for all $f \in Y$. This means that $\mathcal{A} \mathbf{t}(f) = \mathbf{t} A f$ for all $f \in Y$. The subspace Y is dense in X and all operators in the last equality are bounded, which implies the claim. ▷

(2) Assume that a sequence of linear operators (A_n) , with A_n an endomorphism of X_n , is uniformly bounded and converges discretely to a bounded endomorphism A of X . Then $DAp(A) = \{f \in Y : Af \in Y\}$. In particular, if $Y = X$ and the equality 6.2.1 (2) holds on some dense subset of X then it holds on the whole X .

◁ Immediate from (1). ▷

6.2.4. From now we assume that, for all n , the spaces X_n are finite-dimensional and all operators A_n and A are normal or hermitian. Then \mathcal{A} is also normal or hermitian. In this case in view of 6.1.10 $\sigma(\mathcal{A}) = \{\circ\lambda : \lambda \in \sigma(A_N)\}$ and the spectrum $\sigma(\mathcal{A})$ of \mathcal{A} is *discrete*; i.e., $\sigma(\mathcal{A}) = \sigma_p(\mathcal{A})$ or, in other words, $\sigma(\mathcal{A})$ consists only of eigenvalues of \mathcal{A} .

We also assume that the sequence (A_n) converges discretely to A . Obviously for normal operators the commutativity of the diagram of 6.2.3 (1) implies the inclusion: $\sigma(A) \subset \sigma(\mathcal{A})$. So, the eigenvectors of \mathcal{A} that correspond to the points of $\sigma(A)$, may be considered as generalized eigenvectors of A . We use below the following notation: $T^{(\lambda)} := \ker(\lambda - T)$ for an arbitrary operator T .

If A is compact operator and $\mathcal{A}(\mathcal{X}) \subset \mathbf{t}(X)$ then

- (1) $\sigma(A) = \sigma(\mathcal{A})$;
- (2) if $\lambda \in \sigma(A)$ and $\lambda \neq 0$ then $\mathcal{A}^{(\lambda)} = \mathbf{t}(A^{(\lambda)})$, and so $\dim(\mathcal{A}^{(\lambda)}) = \dim(A^{(\lambda)})$.

◁ We have to show that $\sigma(\mathcal{A}) \supset \sigma(A)$ and each eigenvector f of \mathcal{A} not belonging to $\ker(\mathcal{A})$ has the shape $f = \mathbf{t}(x)$ where x is an eigenvector of A with the same eigenvalue as f .

Indeed, take $\lambda \in \sigma(\mathcal{A})$. It is possible to assume that $\lambda \neq 0$. Otherwise $\lambda \in \sigma(A)$ since A is compact and normal. Since $\sigma(\mathcal{A}) = \sigma_p(\mathcal{A})$, there is some f in \mathcal{X} satisfying $\mathcal{A}f = \lambda f$. By hypothesis $\mathcal{A}f \in \mathbf{t}(X)$, and since $\lambda \neq 0$ we have $f \in \mathbf{t}(X)$.

Hence, there is a unique $x \in X$ satisfying $f = \mathbf{t}(x)$. Since the diagram of 6.2.3 (1) commutes, $\mathbf{t}(Ax) = \mathcal{A}(\mathbf{t}(x)) = \mathcal{A}f = \lambda f$. This implies that $\mathbf{t}(\lambda x) = \lambda f$. Furthermore, since \mathbf{t} is injective, we infer that x is an eigenvector of A with eigenvalue λ . ▷

6.2.5. Under the hypotheses of 6.2.3 if A and A_N are hermitian then

- (1) $\sigma_\lambda := \{\nu \in \sigma(A_N) : \nu \approx \lambda\}$ is a finite set for each nonzero $\lambda \in \sigma(\mathcal{A})$; i.e., $\sigma_\lambda = \{\nu_1, \dots, \nu_k\}$ where $k \in \mathbb{N}$;
- (2) The dimension $m_j := \dim(A_N^{(\nu_j)})$ is finite for all $j \leq m$ and $\sum_{j=1}^k m_j = \dim(\mathcal{A}^{(\lambda)})$;
- (3) $(A_N^{(\nu_1)} \oplus \dots \oplus A_N^{(\nu_m)})^\# = \mathcal{A}^{(\lambda)}$;

- (4) If the family $(x_1^j, \dots, x_{m_j}^j)$ is a Hilbert basis for $A_N^{(\nu_j)}$ for all $j := 1, \dots, k$ then $((x_1^1)^\#, \dots, (x_{m_1}^1)^\#, \dots, (x_1^k)^\#, \dots, (x_{m_k}^k)^\#)$ is a Hilbert basis for $\mathcal{A}^{(\lambda)}$;
- (5) If $A^{(\lambda)} \subset \text{Dap}(A)$ then there is a Hilbert basis $(y_1^1, \dots, y_{m_1}^1, \dots, y_1^k, \dots, y_{m_k}^k)$ for $A^{(\lambda)}$ such that $T_N y_l^j \approx x_l^j$ for $j := 1, \dots, k; l := 1, \dots, m_j$.

◁ It is easy to see that $x^\# \in \mathcal{A}^{(\lambda)}$ whenever $\nu \approx \lambda$ and $x \in A_N^{(\nu)}$. In case $\nu_1, \dots, \nu_k \in \sigma_\lambda$ are pairwise distinct, all $x_1 \in A_N^{(\nu_1)}, \dots, x_k \in A_N^{(\nu_k)}$ are pairwise orthogonal and so $x_1^\#, \dots, x_k^\#$ are pairwise orthogonal too. This proves that $k \leq \dim(\mathcal{A}^{(\lambda)})$. The reverse inequality was proven in 6.1.6 (3).

The implications (2) \rightarrow (3) and (3) \rightarrow (4) are obvious, while (5) follows from the definition of \mathbf{t} . ▷

6.2.6. A sequence of operators (A_n) is *quasicompact* provided that $\mathcal{A}(\mathcal{X}) \subset \mathbf{t}(X)$ for all unlimited N (note that (A_n) is not assumed convergent discretely to A).

The reason behind this definition is as follows: Assume for a moment that $((X_n, T_n))_{n \in \mathbb{N}}$ is a strong discrete approximant (see 6.2.1). Then the condition 6.2.1 (2) means that, for all unlimited N , the image $A_N x$ of each limited x is infinitely close to $T_N y$ for some standard $y \in X$. Observe now that the infinitesimal test for an operator to be compact reads (cf. 4.3.6): *T is a compact operator if and only if the image of every limited element under T is nearstandard.*

We now give one simple sufficient condition of quasicompactness of a sequence of operators (A_n) which holds for all Banach spaces X and X_n with $n \in \mathbb{N}$ (cf. [405]). This rests on requiring the following property of discrete approximation which we will often use in the sequel:

$$(1) \sup_{n \in \mathbb{N}} \sup_{\|z\|=1} (\inf\{\|x\|_n : T_n x = z\}) < +\infty.$$

(2) Assume that a sequence (A_n) converges discretely to a compact operator A and, moreover, this convergence is uniform. Assume further that the discrete approximation $((X_n, T_n))_{n \in \mathbb{N}}$ enjoys the property (1). Then (A_n) is a quasicompact sequence.

◁ Take $\xi \in \mathcal{X}$. There is a limited element x in X_N satisfying $\xi = x^\#$. Since every T_n is surjective, by transfer, $x = T_N f$ for some $f \in {}^*X$.

From the hypothesis about $((X_n, T_n))_{n \in \mathbb{N}}$ it follows that f may be chosen limited. By uniform convergence $\|A_N T_N - T_N^* A\|_N \approx 0$ and so $A_N x = A_N T_N f \approx T_N^* A f$.

Since A is compact and f is limited, there is a standard element $h \in X$ satisfying ${}^*A f \approx h$. By 6.2.2, $\|T_N\|_N$ is limited and so $T_N^* A f \approx T_N h$, i.e., $A_N x \approx T_N h$ and $\mathcal{A}(\xi) = \mathbf{t}(h)$. ▷

Below in 7.6 and 7.7 we will use strong discrete convergence in the situation [86] in which there are isometric embeddings $\iota_n : X_n \rightarrow X$ and $T_n := \iota_n^{-1} \circ p_n$, where $p_n : X \rightarrow \iota_n(X_n)$ is the corresponding orthoprojection. It is easy to see that the discrete approximation in this case enjoys the above extra requirement.

Unfortunately, uniform convergence is a rare phenomenon. Therefore, to prove quasicompactness for (A_n) is not an easy matter in many interesting cases. We must first of all find the conditions for $x \in X_N$ to be infinitesimally close to some element of the shape $T_N y$ with some standard $y \in X$. In 7.6.15 we give one test useful for the approach we pursue to approximating Schrödinger type operators.

6.2.7. The definition of quasicompactness acquires a rather natural standard version since condition of the definition of 6.2.6 holds for every unlimited integer N . Then the routine infinitesimal arguments will allow us to prove the following proposition which holds for arbitrary parameters.

We denote the unit ball of X_N with center the origin by B_N , keeping the symbol $B_N(\varepsilon, x)$ for the ball in X_N with center x and radius ε .

If $((X_n, T_n))_{n \in \mathbb{N}}$ is a strong discrete approximant then (A_n) is a quasicompact sequence of operators if and only if

$$(\forall \varepsilon > 0)(\exists^{\text{fin}} B \subset X)(\exists n_0)(\forall N > n_0) \left((A_N(B_N)) \subset \bigcup_{y \in B} B_N(\varepsilon, T_N y) \right).$$

$\triangleleft \leftarrow$: Take an arbitrary standard $\varepsilon > 0$. By transfer, we have the following inclusion:

$$A_N(B_N) \subset \bigcup_{y \in B} B_N(\varepsilon, T_N y)$$

for every unlimited natural N . This implies for $x \in B_N$ that to each standard $n \in \mathbb{N}$ there is a standard $y_n \in X$ satisfying $\|A_N x - T_N y_n\|_N < n^{-1}$. Since $\|T_N y_n - T_N y_m\|_N \approx \|y_n - y_m\|$ for all standard n and m ; therefore, $(y_n)_{n \in \mathbb{N}}$ is a Cauchy sequence in X and so $(y_n)_{n \in \mathbb{N}}$ converges to some standard element $y \in X$. Now it is obvious that $\|A_N x - T_N y\| \approx 0$.

\rightarrow : Suppose the contrary. By negation,

$$(\exists \varepsilon > 0)(\forall^{\text{fin}} B \subset X)(\forall n_0)(\exists N > n_0) \left((A_N(B_N)) \not\subset \bigcup_{y \in B} B_N(\varepsilon, T_N y) \right).$$

Take a standard $\varepsilon_0 > 0$ satisfying the last formula. Consider a hyperfinite set $B \subset {}^*X$ such that $X \subset B$.

By transfer, from the above formula it follows that there is an unlimited N satisfying

$$A_N(B_N) \not\subset \bigcup_{y \in B} B_N(\varepsilon, T_N y).$$

Thus, there is some x in B_N such that the distance from x to each element of the shape $T_N y$ with a standard y is greater than or equal to the standard ε_0 . This shows that the condition of the definition of quasicompactness in 6.2.6 fails for this number N . \triangleright

6.2.8. Theorem. *Let A be a compact hermitian endomorphism of a Hilbert space X . Assume that $((X_n, T_n))_{n \in \mathbb{N}}$ is a discrete approximant to X , and $(A_n)_{n \in \mathbb{N}}$ is a quasicompact sequence convergent discretely to A , with A_n a hermitian endomorphism of X_n for all n . Then*

- (1) *The spectrum $\sigma(A)$ of A coincides with the set of nonisolated limit points of $\bigcup_n \sigma(A_n)$;*
- (2) *If $0 \neq \lambda \in \sigma(A)$ and J is a neighborhood of λ with no points of $\sigma(A)$ distinct from λ then λ is the sole nonisolated limit point of $J \cap \bigcup_n \sigma(A_n)$;*
- (3) *If in the context of (2)*

$$M_n^\lambda := \sum_{\nu \in \sigma(A_n) \cap J} A_n^{(\nu)},$$

then $\dim(M_n^\lambda) = \dim(A^{(\lambda)}) = s$ for all sufficiently large n and there is a sequence of orthonormal bases (f_1^n, \dots, f_s^n) for M_n^λ , convergent discretely to an orthonormal basis (f_1, \dots, f_s) for $A^{(\lambda)}$.

\triangleleft This is a standard rephrasing of 6.2.4 and 6.2.5. \triangleright

Certainly, our arguments fail in the case of unbounded selfadjoint operator A . Indeed, if (A_n) converges discretely to A then the norm of A_N is unlimited and we face problems even in defining $A_N^\#$ (see [250]). However, we may obviate the obstacles by slightly modifying the results on strong resolvent convergence (cf., for example, [409, Theorem VIII.19] and also [117, Theorem 5.7.6]). It is worth noting also that all facts of the theory of unbounded linear operators we use in the sequel may be found in [409, Chapter VIII].

If $\lambda \notin \sigma(A)$ then we let $R_\lambda(A) := (\lambda - A)^{-1} := (\lambda 1 - A)^{-1}$ denote the resolvent of A at λ , with 1 standing as usual for the identity operator I on X , the unity of the endomorphism algebra of X .

6.2.9. *If A is selfadjoint and the approximation domain $D_{Ap}(A)$ of A by (A_n) is an essential domain of A then $(R_\lambda(A_n))$ converges discretely to $R_\lambda(A)$ for all λ such that $\lambda \notin \text{cl}(\sigma(A) \cup \bigcup_n \sigma(A_n))$.*

◁ Start with proving the claim for $\lambda := \pm i$ which obviously meets the assumptions. Consider only the case of $\lambda := -i$ since the case of $\lambda := i$ is settled similarly. By definition, it suffices to prove 6.2.1 (2) with $R_{-i}(A)$ and $R_{-i}(A_n)$ and some dense subset $Y \subset X$.

Put $Y := \{(A + i)\varphi : \varphi \in DA_p(A)\}$. Then Y is dense in X because $DA_p(A)$ is an essential domain of A . Take an unlimited natural N and note

$$((A + i)^{-1} - (A_N + i)^{-1})(A + i)\varphi = (A_N + i)^{-1}(A_N - A)\varphi.$$

Since $\varphi \in DA_p(A)$; therefore, $(A_N - A)\varphi \approx 0$.

Given a subset B of the reals \mathbb{R} and $\lambda \in \mathbb{C}$, denote the distance from λ to B on the plane \mathbb{C} by $\rho(\lambda, B)$. Assign $S := \text{cl}(\sigma(A) \cup \bigcup_n \sigma(A_n))$. If λ satisfies the hypotheses then $\rho(\lambda, S) > 0$. By transfer,

$$\|(A_N + i)^{-1}\| = \rho(\lambda, \sigma(A_N))^{-1} \leq \rho(\lambda, S)^{-1} \ll +\infty.$$

Thus, $R_{-i}(A_N)(A + i)\varphi \approx R_{-i}(A)(A + i)\varphi$. This proves that $R_{-i}(A_n)$ converges discretely to $R_{-i}(A)$.

We prove now that if the claim is valid for some λ_0 such that $\lambda \notin S$ and if $|\lambda - \lambda_0| < \rho(\lambda_0, S)$ then the claim holds for this λ . This will clear do since each $\lambda \notin S$ can be connected either with i or with $-i$ by a smooth curve lying entirely in $\mathbb{C} - S$. So we may reach λ from i or $-i$ through finitely many circles of radius less than $\rho(\lambda, S)$.

The functions $R_\lambda(A_n)$ ($n \in \mathbb{N}$) and $R_\lambda(A)$ are analytic in the open disk $\{\lambda \in \mathbb{C} : |\lambda - \lambda_0| < \rho(\lambda_0, S)\}$, and expand in the following uniformly convergent series:

$$(1) R_\lambda(A) = \sum_{m=0}^{\infty} (\lambda - \lambda_0)^m R_{\lambda_0}^{m+1}(A),$$

$$(2) R_\lambda(A_n) = \sum_{m=0}^{\infty} (\lambda - \lambda_0)^m R_{\lambda_0}^{m+1}(A_n).$$

Demonstrate that $T_N R_\lambda(A)f \approx R_\lambda(A_N)T_N f$ for all unlimited N and standard $f \in Y$. Since this holds for λ_0 ; therefore,

$$(3) \sum_{m=0}^k (\lambda - \lambda_0)^m T_N R_{\lambda_0}^{m+1}(A)f \approx \sum_{m=0}^k (\lambda - \lambda_0)^m R_{\lambda_0}^{m+1}(A_N)T_N f$$

for all $k \in \mathbb{N}$.

Take an arbitrary standard $\varepsilon > 0$. Then by (1) and (2) there is some n_0 satisfying

$$\left\| R_\lambda(A)f - \sum_{m=0}^k (\lambda - \lambda_0)^m R_{\lambda_0}^{m+1}(A)f \right\| < \varepsilon$$

for $k > n_0$. Denote the function on the left part of the last inequality by h . Since h is standard, $\|T_N h\| \approx \|h\|$. This implies that

$$(4) \left\| T_N R_\lambda(A) f - \sum_{m=0}^k (\lambda - \lambda_0)^m T_N R_{\lambda_0}^{m+1}(A) f \right\|_N \leq \varepsilon.$$

Show now that the convergence of the series in (2) is uniform in n . To this end, note that

$$\|R_{\lambda_0}(A_n)\| = \max_{\nu} \{\nu : \nu \in \sigma(A_n)\} = \rho(\lambda_0, \sigma(A_n))^{-1} \leq \rho(\lambda_0, S)^{-1}.$$

Thus, $q := |\lambda - \lambda_0| \|R_{\lambda_0}(A_n)\| < 1$, implying that

$$\left\| \sum_{m=k+1}^{\infty} (\lambda - \lambda_0)^m R_{\lambda_0}^{m+1}(A_n) \right\| \leq \frac{\rho(\lambda_0, S)^{-1} q^{k+1}}{1 - q} \rightarrow 0$$

for $k \rightarrow \infty$. Now, since $\|T_N f\|_N$ is limited ($T_N f \approx f$); there is some n_1 satisfying

$$(5) \left\| R_\lambda(A_n) T_N f - \sum_{m=0}^k (\lambda - \lambda_0)^m R_{\lambda_0}^{m+1}(A_n) T_N f \right\|_N < \varepsilon$$

for $k > n_1$.

Taking a standard $k > \max\{n_0, n_1\}$, infer by (3), (4), and (5) that

$$\|T_N R_\lambda(A) f - R_\lambda(A_N) T_N f\|_N \leq 2\varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, the proof is complete. \triangleright

6.2.10. The following proposition is a simple corollary of the above fact.

If in the conditions of 6.2.9 the resolvents of selfadjoint operators A_n are compact and $(R_\lambda(A))$ is a quasicompact sequence for some real $\lambda \notin \text{cl}(\sigma(A) \cup \bigcup_n \sigma(A_n))$ then 6.2.8 (1)–(3) are valid.

6.2.11. If $((X_n, T_n))_{n \in \mathbb{N}}$ is a strong discrete approximant satisfying 6.2.6 (1) then by 6.2.6 (2) the uniform convergence of the sequence of resolvents implies that this is a quasicompact sequence. The following proposition (cf. [409, Theorem 8.25]) provides a sufficient condition for uniform convergence of resolvents.

Suppose that in the conditions of 6.2.9 the discrete approximant $((X_n, T_n))_{n \in \mathbb{N}}$ is strong and satisfies 6.2.6 (1) and for some essential domain $D \subset DAp(A)$ of A the following condition holds

$$\lim_{n \rightarrow \infty} \sup_{\varphi \in D, \|\varphi\|_A = 1} \|(T_n A - A_n T_n) \varphi\|_n = 0,$$

where $\|\varphi\|_A := \|A\varphi\| + \|\varphi\|$. Then the convergence of $(R_\lambda(A_n))$ to $R_\lambda(A)$ is uniform.

6.2.12. COMMENTS.

(1) The origin of this section is the article [5] by Albeverio, Gordon, and Khrennikov. The definitions of 6.2.1 specify the concept of discrete convergence which stems from the works of Stummel [468, 469]. Some aspects of this notion are discussed in the book [410]. We use the notion of *discrete compactness* which for our particular case amounts to discrete compactness as introduced in [410, 7.3].

(2) Discrete approximation in the case of $Y = X$ is also used in [86] wherein X_n are some spaces of grid functions embedded into $L_2(\mathbb{R}^n)$ as the spaces of step functions and the orthoprojections onto the latter serve as T_n . Note that in this case $T_n f$ differs from the table of f at the knots of the grid even for a continuous function f .

However, if f is smooth then it is easy to see that the difference vanishes between $T_n f$ and the table of f as n tends to infinity. The possibility of considering A_n as an endomorphism of X (in this case $X = L_2(\mathbb{R}^n)$) simplifies the study of the limit behavior of the spectra of approximants. The key tool for this is the concept of *uniform compactness* which means that $\bigcup_{n \in \mathbb{N}} A_n(U)$, with U the unit ball of $L_2(\mathbb{R}^n)$, is relatively compact (see [86]). This concept was suggested in [15].

(3) We consider here only selfadjoint operators with discrete spectra and compact resolvents but it seems that our approach may be useful in other cases. Some general results on the limit behavior of the spectra of approximants A_n converging discretely to some operator A in the general Banach spaces but in the case of $Y = X$ (see the definitions of 6.2.1) were obtained by Rübiger and Wolff [405, 528] also by infinitesimal methods. Gordon in [146] pursues the same approach in the special case of selfadjoint integral operators on compact groups.

(4) The proof of 6.2.7 involves $\text{Card}(X)^+$ -saturation, i.e., the properly-stated concurrence principle. Clearly, in case X is separable, it suffices to use ω_1 -saturation of the nonstandard universe.

(5) It is interesting to compare the quasicompactness property of 6.2.6 with the compactness property of the sequence (A_n) as in [86].

As was mentioned, the environment of [86] assumes some isometric embeddings $\iota_n : X_n \rightarrow X$ and $T_n := \iota_n^{-1} \circ p_n$, with $p_n : X \rightarrow \iota_n(X_n)$ the orthoprojection from X onto $\iota_n(X_n)$. In this event we may identify A_n with the endomorphism $\bar{A}_n := \iota_n A_n T_n$ of X .

By [86] (A_n) is a *compact sequence*, provided that $\bigcup_n \bar{A}_n(B)$, with B the unit ball of X , is relatively compact. It is easy to see that if (A_n) is a compact sequence in the sense of [86] then (A_n) is quasicompact. Indeed, if $x \in B_N$ for all infinite $N \in {}^*\mathbb{N}$ then $\iota_N(x) \in B$ and $\bar{A}_N \iota_N x = \iota_N A_N x \approx y$ for some standard $y \in X$ by the compactness property in the sense of [86] and Theorem 4.3.6.

6.3. Loeb Measure

The construction of Loeb measure is one of the most noticeable achievements of infinitesimal analysis which gave rise to applications in many sections of functional analysis, probability, and stochastic modeling; see [3, 69]. We now present a few results about the structure of Loeb measure.

6.3.1. Let (X, \mathcal{A}, ν) be an internal measure space with ν a finitely additive measure; more exactly, assume that \mathcal{A} is an internal algebra of subsets of an internal set X and $\nu : \mathcal{A} \rightarrow {}^*\mathbb{R}$ is an internal finitely additive function on \mathcal{A} .

This implies in particular that $\mathcal{A} \subset {}^*\mathcal{P}(X)$. Moreover, if $\{A_1, \dots, A_\Omega\}$ is a hyperfinite set of members of \mathcal{A} , with $\Omega \in {}^*\mathbb{N}$, then $\bigcup_{k=1}^\Omega A_k \in \mathcal{A}$; if A_k are pairwise disjoint then

$$\nu \left(\bigcup_{k=1}^\Omega A_k \right) = \sum_{k=1}^\Omega \nu(A_k).$$

The union of a hyperfinite family of sets is defined in much the same way as the sum of a hyperfinite set in 6.1.3. If $f : \mathbb{N} \rightarrow \mathcal{P}(X)$ is a sequence in $\mathcal{P}(X)$ then we define the new sequence $g : \mathbb{N} \rightarrow \mathcal{P}(X)$ by recursion:

$$\text{Seq}(f) \wedge \text{Seq}(g) \wedge f(0) = g(0) \wedge (\forall k \in \mathbb{N})(g(k+1) = g(k) \cup f(k+1)).$$

Abbreviate this formula to $\bigcup(f, g)$, take a hyperfinite set $\mathcal{A}_0 \subset \mathcal{A}$, and put $\Omega := |\mathcal{A}_0|$. Proceed as follows: Distinguish some bijection $f : \{0, \dots, \Omega - 1\} \rightarrow \mathcal{A}_0$ and extend it to the internal sequence $f : {}^*\mathbb{N} \rightarrow \mathcal{A}$ by letting f be zero for $n > \Omega - 1$. We now define the sequence $g : {}^*\mathbb{N} \rightarrow \mathcal{A}$ by the rule $\bigcup(f, g)$. By transfer, $g(\Omega - 1)$ is independent of the choice of f . Therefore, we may soundly assign $\bigcup_{A \in \mathcal{A}_0} A := g(\Omega - 1)$. Again by transfer, the so-defined union of a hyperfinite family enjoys all properties of the conventional finite union.

Consider the external function

$${}^\circ\nu : A \mapsto {}^\circ(\nu(A)) \in \mathbb{R}^\bullet \quad (A \in \mathcal{A}),$$

where ${}^\circ(\nu(A))$ is as usual the standard part of $\nu(A)$ whenever $\nu(A)$ is limited and ${}^\circ(\nu(A)) = +\infty$ otherwise. It is easy to see that ${}^\circ\nu$ is a finitely additive function.

6.3.2. The Loeb measure for ν appears as a unique countably additive extension of ${}^\circ\nu$ to the external σ -algebra $\sigma(\mathcal{A})$ generated by \mathcal{A} . The existence of this extension, as will be shown in 6.3.4, ensues from the Lebesgue–Carathéodory Extension Theorem, whereas the uniqueness of this extension requires a few auxiliary facts of a technical nature.

(1) Let \mathcal{A}_0 be a countable subalgebra of \mathcal{A} and let \mathcal{B}_0 be a complete algebra of subsets of X (i.e., a complete subalgebra of $\mathcal{P}(X)$) generated by \mathcal{A}_0 . If $S \subset X$ and for every $A \in \mathcal{A}_0$ either $S \subset A$ or $S \cap A = \emptyset$ then for every $B \in \mathcal{B}_0$ either $S \subset B$ or $S \cap B = \emptyset$.

◁ Call a subset P of X marked provided that $P := \bigcap_{k=1}^{\infty} B_k$ where $(B_n) \subset \mathcal{A}_0$ and for every member A of \mathcal{A}_0 either A or the complement of A to X coincides with one of the sets B_k . Let \mathcal{P} stand for the set of all marked subsets of X . Clearly, the members of \mathcal{P} are pairwise disjoint and $\bigcup \mathcal{P} = X$; i.e., \mathcal{P} is a *partition* of X . By the definition of marked set, if $P \in \mathcal{P}$ and $A \in \mathcal{A}_0$ then either $P \subset A$ or $P \cap A = \emptyset$. Hence, $A = \bigcup \{P \in \mathcal{P} : P \subset A\}$ for all $A \in \mathcal{A}_0$. Note also that \mathcal{B}_0 consists exactly of the sets $B \subset X$ of the shape $B := \bigcup \mathcal{P}'$, with $\mathcal{P}' \subset \mathcal{P}$.

We now take a set $S \subset X$ such that for every $A \in \mathcal{A}_0$ either $S \subset A$ or $S \cap A = \emptyset$. Then the sequence (B_k) of all elements of \mathcal{A}_0 including S is such that $P := \bigcap B_k$ belongs to \mathcal{P} . Since $S \subset P$; therefore, if $B \in \mathcal{B}_0$ then the above shows that only two cases are possible: either $P \subset B$ implying $S \subset B$ or $P \cap B = \emptyset$ implying $S \cap B = \emptyset$. ▷

Let $c(\mathcal{A})$ stand for the collection of all sets $S \subset X$ satisfying the following condition: There is a countable subalgebra \mathcal{A}_0 of \mathcal{A} such that S belongs to a complete subalgebra of subsets of X generated \mathcal{A}_0 (i.e., in a complete subalgebra of the powerset $\mathcal{P}(X)$). In this event, we say that S is *generated* by \mathcal{A}_0 . The following is immediate from the above definitions.

(2) The set $c(\mathcal{A})$ is a σ -algebra and $\sigma(\mathcal{A}) \subset c(\mathcal{A})$.

6.3.3. If $S \in c(\mathcal{A})$ then the following dilemma is open:

- (1) There is some $A \in \mathcal{A}$ such that $A \subset S$ and $\nu(A)$ is an unlimited real;
- (2) There is a sequence $(A_k)_{k \in \mathbb{N}} \subset \mathcal{A}$ such that $S \subset \bigcup_k A_k$ and $\nu(A_k)$ is a limited real for all $k \in \mathbb{N}$.

◁ Let $S \in c(\mathcal{A}_0)$ be a set generated by the countable subalgebra \mathcal{A}_0 of \mathcal{A} . Put $\mathcal{A}'_0 := \{A \in \mathcal{A}_0 : |\nu(A)| \ll +\infty\}$. If $S \subset \bigcup \mathcal{A}'_0$ then (2) holds. In the opposite case, take $p \in S - \bigcup \mathcal{A}'_0$ and consider the countable set $\mathcal{A}''_0 := \{A \in \mathcal{A}_0 : p \in A\}$. Note that \mathcal{A}''_0 has the finite intersection property and consists of sets of unlimited measure. Put $A(n, B) := \{A \in \mathcal{A} : p \in A \subset B, \nu(A) \geq n\}$. By the properties of \mathcal{A}''_0 , the set $\{A(n, B) : n \in \mathbb{N}, B \in \mathcal{A}''_0\}$ has the finite intersection property. By saturation, there is some A in \mathcal{A} satisfying $A \in A(n, B)$ for all $n \in \mathbb{N}$ and $B \in \mathcal{A}''_0$. Hence, $\nu(A)$ is unlimited, $p \in A$, and $A \subset B$ for all $B \in \mathcal{A}''_0$.

Assume that some C in \mathcal{A}_0 does not included A . Then either $p \notin C$ or $p \in X - C$, implying that $X - C \in \mathcal{A}''_0$ and so $A \subset X - C$. Thus, if $C \in \mathcal{A}_0$ then either $A \subset C$ or $A \cap C = \emptyset$. Since S is generated by \mathcal{A}_0 ; therefore, either $A \subset S$ or $A \cap S = \emptyset$ by 6.3.2 (1). Since $p \in A \cap S$, it follows that $A \subset S$, implying (1).

Assume that (1) and (2) hold simultaneously. In this event $A \subset \bigcup_k A_k$ and $\nu(A)$ is unlimited whereas $\nu(A_k)$ are all limited. By saturation, $A \subset A_1 \cup \cdots \cup A_n$ for some $n \in \mathbb{N}$. This is however impossible implying $\nu(A) \leq \nu(A_1) + \cdots + \nu(A_n)$ which is a contradiction. \triangleright

6.3.4. Theorem. *A finitely additive measure $\circ\nu : \mathcal{A} \rightarrow \mathbb{R}^\bullet$ admits a unique countably additive extension λ to the external σ -algebra $\sigma(\mathcal{A})$ generated by \mathcal{A} . Moreover:*

- (1) $\lambda(B) = \inf\{\circ\nu(A) : B \subset A, A \in \mathcal{A}\} \quad (B \in \sigma(\mathcal{A}));$
- (2) *If $\lambda(B) < +\infty$ for some $B \in \sigma(\mathcal{A})$ then $\lambda(B) = \sup\{\circ\nu(A) : A \subset B, A \in \mathcal{A}\};$*
- (3) *If $\lambda(B) < +\infty$ for some $B \in \sigma(\mathcal{A})$ then there is $A \in \mathcal{A}$ satisfying $\lambda(A \Delta B) = 0;$*
- (4) *To an arbitrary $B \in \sigma(\mathcal{A})$ either there is $A \in \mathcal{A}$ such that $A \subset B$ and $\circ\nu(A) = +\infty$, or there is a sequence $(A_n)_{n \in \mathbb{N}}$ of sets in \mathcal{A} such that $B \subset \bigcup_{n \in \mathbb{N}} A_n$ and $\circ\nu(A_n) < +\infty$ for all $n \in \mathbb{N}$.*

\triangleleft The existence of λ follows from the Lebesgue–Carathéodory Extension Theorem. The validation of the hypotheses of this theorem is trivial. Indeed, take an increasing sequence of sets $(A_k)_{k \in \mathbb{N}}$ in \mathcal{A} and suppose that $A := \bigcup_k A_k$ belongs to \mathcal{A} . By saturation, $A = A_m$ for some $m \in {}^*\mathbb{N}$ and so $\circ\nu(A_k) \rightarrow \circ\nu(A)$.

We now prove (1)–(3) and the claim of uniqueness. Take $B \in \sigma(\mathcal{A})$. The Lebesgue–Carathéodory Extension Theorem guarantees in particular that

$$\lambda(B) = \inf \left\{ \sum_{k=1}^{\infty} \circ\nu(A_k) : A_k \in \mathcal{A} (k \in \mathbb{N}), B \subset \bigcup_{k=1}^{\infty} A_k \right\}$$

for all $B \in \sigma(\mathcal{A})$. Consequently, to each $0 < \varepsilon \in \mathbb{R}$ there is a sequence of internal sets $(A_k)_{k \in \mathbb{N}}$, satisfying $B \subset \bigcup_{k=1}^{\infty} A_k$ and $\sum_{k=1}^{\infty} \circ\nu(A_k) < \lambda(B) + \varepsilon/2$. Since $\nu(A_k) < \circ\nu(A_k) + \varepsilon/2^{k+1}$ for all $k \in \mathbb{N}$; therefore, given $n \in \mathbb{N}$, we may write

$$\nu \left(\bigcup_{k=1}^n A_k \right) \leq \sum_{k=1}^n \nu(A_k) \leq \sum_{k=1}^{\infty} \circ\nu(A_k) + \sum_{k=1}^{\infty} \varepsilon/2^{k+1} < \lambda(B) + \varepsilon.$$

Extend $(A_k)_{k \in \mathbb{N}}$ to some internal sequence $(A_k)_{k \in {}^*\mathbb{N}}$ by 3.5.11 (1). Consider the internal set $\{n \in {}^*\mathbb{N} : \nu(\bigcup_{k=1}^n A_k) < \lambda(B) + \varepsilon\}$. Since this set contains all standard naturals, it also contains some unlimited natural Ω by overflow. Put $A_\Omega := \bigcup_{k=1}^{\Omega} A_k$. By definition, $B \subset A_\Omega$ and $\nu(A_\Omega) < \lambda(B) + \varepsilon$ and so $\circ\nu(A_\Omega) \leq \lambda(B) + \varepsilon$. This proves (1).

Assume that $\lambda(B) < +\infty$. By what was proven there is an internal set $C \in \mathcal{A}$ such that $B \subset C$ and $\nu(C)$ is finite. Applying (1) to $C - B$, infer (2). Furthermore, let $(A_k)_{k \in \mathbb{N}}$ be an increasing sequence in \mathcal{A} satisfying $A_k \subset B$ and

$|\nu(A_k) - \lambda(B)| < 1/k$. By extension, consider this sequence as a part of an internal increasing sequence $(A_k)_{k \in {}^*\mathbb{N}}$ with the same property. By overflow, there is an unlimited hypernatural $\Omega \in {}^*\mathbb{N}$ satisfying $|\nu(A_\Omega) - \lambda(B)| < 1/\Omega$. Then $\lambda(B) = \circ\nu(A_\Omega)$ and it is easy to see that $\lambda(A_\Omega \triangle B) = 0$. Infer (4) now from 6.3.3.

We are left with proving uniqueness. Suppose that λ_1 and λ_2 are two σ -additive extensions of the measure $\circ\nu$ to $\sigma(\mathcal{A})$. Since $\sigma(\mathcal{A}) \subset c(A)$, it is possible to apply 6.3.3 to $S \in \sigma(\mathcal{A})$. In the case of 6.3.3 (1), there is some $A \in \mathcal{A}$ satisfying $A \subset S$ and $\circ\nu(A) = \infty$, implying that $\lambda_j(S) = \infty$ ($j := 1, 2$). In the case of 6.3.3 (2), there is a sequence $(A_k)_{k \in \mathbb{N}}$ in \mathcal{A} such that $S \subset \bigcup_k A_k$ and $\nu(A_k)$ is limited for all $k \in \mathbb{N}$. Without loss of generality, we may assume that this sequence increases. By 6.3.4 (1)

$$\lambda_j(S \cap A_k) = \inf\{\circ\nu(A) : A \in \mathcal{A}, S \cap A_k \subset A \subset A_k\} \quad (j := 1, 2).$$

Hence we infer in particular that $\lambda_1(S \cap A_k) = \lambda_2(S \cap A_k)$ for $k \in \mathbb{N}$. Since $S = \bigcup_k (S \cap A_k)$ and $(S \cap A_k)_{k \in \mathbb{N}}$ is an increasing sequence; therefore, $\lambda_1(S) = \lambda_2(S)$. Thus, λ_1 and λ_2 coincide on $\sigma(\mathcal{A})$. \triangleright

6.3.5. Let $S(\mathcal{A})$ stand for a completion of $\sigma(\mathcal{A})$ with respect to λ , and let ν_L stand for the extension of λ to $S(\mathcal{A})$. We may show that, in case $\nu_L(X) < +\infty$, the membership $B \in S(\mathcal{A})$ holds if and only if

$$\sup\{\circ\nu(A) : A \subset B, A \in \mathcal{A}\} = \inf\{\circ\nu(A) : B \subset A, A \in \mathcal{A}\} = \nu_L(B).$$

The triple $(X, S(\mathcal{A}), \nu_L)$, presenting a measure space with the σ -additive measure ν_L , is the *Loeb measure space* (for (X, \mathcal{A}, ν)); and ν_L , the *Loeb measure* (for ν).

A function $f : X \rightarrow {}^*\overline{\mathbb{R}}$ is *Loeb-measurable* provided that f is measurable with respect to the σ -algebra $S(\mathcal{A})$. An internal function $F : X \rightarrow {}^*\mathbb{R}$ is \mathcal{A} -measurable if $\{x \in X : F(x) \leq t\} \in \mathcal{A}$ for all $t \in {}^*\mathbb{R}$. An internal \mathcal{A} -measurable function $F : X \rightarrow {}^*\mathbb{R}$ is a *lifting* of a function $f : X \rightarrow {}^*\overline{\mathbb{R}}$ whenever $f(x) = \circ F(x)$ for ν_L -almost all $x \in X$.

6.3.6. Theorem. *A function $f : X \rightarrow {}^*\overline{\mathbb{R}}$ is Loeb-measurable if and only if f has a lifting.*

$\triangleleft \leftarrow$: Take an \mathcal{A} -measurable internal function $F : X \rightarrow {}^*\mathbb{R}$. Given arbitrary standard $r \in \mathbb{R}$, note that

$$\{x \in X : \circ F(x) \leq r\} = \bigcap_{k \in \mathbb{N}} \{x \in X : F(x) \leq r + 1/k\} \in \sigma(\mathcal{A}),$$

and so the function $\circ F$ is Loeb-measurable. If $f(x) = \circ F(x)$ ν_L -almost everywhere then f is Loeb-measurable too.

→: Assume that f is Loeb-measurable. Consider an arbitrary numbering $(q_k)_{k \in \mathbb{N}}$ of the *rational*s \mathbb{Q} ; in symbols, $\mathbb{Q} = \{q_k : k \in \mathbb{N}\}$. Put $B_k := \{x \in X : f(x) \leq q_k\}$. Choose internal sets $A_k \in \mathcal{A}$ so that $\nu_L(A_k \Delta B_k) = 0$ and $A_k \subset A_l$ for $q_k \leq q_l$. By extension and overflow, there is an unlimited natural Ω such that $A_k \in \mathcal{A}$ and if $q_k \leq q_l$ then $A_k \subset A_l$ for all $k, l \leq \Omega$. Define some internal \mathcal{A} -measurable function $F : X \rightarrow {}^*\mathbb{R}$ with range $\{q_1, \dots, q_\Omega\}$ by the condition that $F(x) \leq q_k$ is equivalent to $x \in A_k$. In more detail, if $\{q_1, \dots, q_\Omega\}$ is renumbered in increasing order $q_{k_1} < q_{k_2} < \dots < q_{k_\Omega}$ then put

$$F(x) := \begin{cases} q_{k_1}, & \text{if } x \in A_{k_1}, \\ q_{k_l}, & \text{if } x \in A_{k_l} - A_{k_{l-1}} \quad (1 < l \leq \Omega), \\ q_{k_{\Omega+1}}, & \text{if } x \notin A_{k_\Omega}. \end{cases}$$

Clearly, if $k \in \mathbb{N}$ then $F(x) \leq q_k$ amounts to $f(x) \leq q_k$ for all $x \notin D := \bigcup_{k \in \mathbb{N}} A_k \Delta B_k$. Since $\nu_L(D) = 0$; therefore, ${}^\circ F(x) = f(x)$ for ν_L -almost all $x \in X$. \triangleright

6.3.7. An internal function F is *simple* if the range $\text{im}(F)$ of F is a hyperfinite set. Inspection of the proof of Theorem 6.3.6 shows that each Loeb-measurable function has a lifting that is a simple function. Obviously, a simple internal function F is \mathcal{A} -measurable if and only if $F^{-1}(\{t\}) \in \mathcal{A}$ for all $t \in {}^*\mathbb{R}$. In this event to F there corresponds the internal integral

$$\int_X F d\nu = \sum_{t \in \text{im}(F)} F(t) \nu(F^{-1}(\{t\})).$$

If $A \in \mathcal{A}$ then, as usual, $\int_A F d\nu = \int_X F \chi_A d\nu$, where χ_A is the characteristic function of a set A .

Put $A_N := \{x \in X : |F(x)| \geq N\}$. An internal simple \mathcal{A} -measurable function $F : X \rightarrow {}^*\mathbb{R}$ is called *\mathcal{S} -integrable* provided $\int_{A_N} F d\nu \approx 0$ for all unlimited N .

It is possible to prove that the \mathcal{S} -integrability of F is equivalent to each of the following conditions:

- (1) $\int_X F d\nu$ is a limited hyperreal number and $\int_A F d\nu \approx 0$ whenever $A \in \mathcal{A}$ and $\nu(A) \approx 0$;
- (2) $\int_X {}^\circ|F| d\nu_L = {}^\circ(\int_X |F| d\nu) < +\infty$.

The next two theorems deal with a finite Loeb measure space: $\nu_L(X) < +\infty$.

6.3.8. Theorem. Let (X, \mathcal{A}, ν) be an internal measure space with finitely additive measure ν and let $(X, \mathcal{S}(\mathcal{A}), \nu_L)$ be the Loeb measure space for (X, \mathcal{A}, ν) . A function $f : X \rightarrow {}^*\mathbb{R}$ is ν_L -integrable if and only if f has an \mathcal{S} -integrable lifting $F : X \rightarrow {}^*\mathbb{R}$. In this case

$$\int_X f d\nu_L = {}^\circ \left(\int_X F d\nu \right).$$

◁ Prove first that if f is bounded, Loeb-measurable and possesses a bounded lifting F then $\int_X f d\nu_L = \circ \int_X F d\nu$. To this end, take a simple Loeb measurable function g with range a finite set $\{r_1, \dots, r_n\}$; moreover, assume that $g \leq f$. By the theorem of 6.3.4, to the set $B_k := g^{-1}(r_k)$ there is some $A_k \in \mathcal{A}$ satisfying $\nu_L(A_k \Delta B_k) = 0$. The function G , equal to r_k on A_k , is a lifting of g . Moreover, $\int_X g d\nu_L = \circ \int_X G d\nu$ since both integrals are finite sums equal to each other.

If a bounded lifting F_1 of f satisfies the inequality then, on the one hand, $\int_X G d\nu_L \leq \int_X F_1 d\nu$ and, on the other hand, $|F(x) - F_1(x)| \leq 1/n$ for ν_L -almost all $x \in X$ and for all standard $n \in \mathbb{N}$. Hence, $\int_X F d\nu \approx \int_X F_1 d\nu$, implying that

$$\int_X g d\nu_L = \circ \left(\int_X G d\nu \right) \leq \circ \left(\int_X F_1 d\nu \right) = \circ \left(\int_X F d\nu \right).$$

Hence, $\int_X f d\nu_L \leq \circ \left(\int_X F d\nu \right)$. The reverse inequality results on replacing f with $-f$ and F with $-F$.

Assume now that f is ν_L -integrable. Without loss of generality, assume further that $f \geq 0$. (The general case will ensue on using the standard representation of F as the difference of its positive and negative parts: $f = f^+ - f^-$.)

Let F' be a lifting of f whose existence follows from Theorem 6.3.6. If $F_n := F' \wedge n$ then the above implies

$$\circ \left(\int_X F_n d\nu \right) = \int_X (f \wedge n) d\nu_L \rightarrow \int_X f d\nu_L.$$

Applying to $(F_n)_{n \in \mathbb{N}}$ the extension and overflow principles, find an unlimited natural Ω satisfying $\circ \left(\int_X F_N d\nu \right) \approx \int_X f d\nu_L \rightarrow \int_X f d\nu_L$ for all unlimited naturals $N \leq \Omega$. The function $F := F_\Omega$ is an \mathcal{S} -integrable lifting of f .

Conversely, assume given an \mathcal{S} -integrable lifting F of f . By underflow, to an arbitrary $0 < \varepsilon \in \mathbb{R}$ there is a standard natural $n \in \mathbb{N}$ such that $\int_{F \geq m} F d\nu \leq \varepsilon$ for each standard $m \geq n$. Applying again the above-indicated assertion for a bounded function, infer that

$$\begin{aligned} \int_X (f \wedge n) d\nu_L &\approx \int_X (F \wedge n) d\nu \leq \int_X F d\nu \\ &\leq \int_X (F \wedge n) d\nu + \varepsilon \approx \int_X (f \wedge n) d\nu_L + \varepsilon. \end{aligned}$$

Hence, $\circ \left(\int_X F d\nu \right)$ is a limited real serving as the limit of the sequence $\int_X (f \wedge n) d\nu_L$ as $n \rightarrow \infty$. This ends the proof. ▷

6.3.9. Theorem. Let (X, \mathcal{A}, ν) be an internal space with ν a finitely additive measure. For each simple internal \mathcal{A} -measurable function the following are equivalent:

- (1) F is \mathcal{S} -measurable;
- (2) $\int_X |F| d\nu < +\infty$ and $\nu(A) \approx 0$ implies $\int_A |F| d\nu \approx 0$ for all $A \in \mathcal{A}$;
- (3) $\int_X \circ |F| d\nu_L = \int_X \circ |F| d\nu$.

◁ The proof proceeds by analogy with no complications. ▷

6.3.10. We suppose now that X is a hyperfinite set, $\mathcal{A} := \mathcal{P}(X)$, and ν is a counting measure multiplied by Δ , i.e., $\nu(A) := \Delta|A|$ for all $A \in \mathcal{A}$, where $|A|$ stands for the size of A and $\Delta \in {}^*\mathbb{R}_+$.

The corresponding Loeb measure space is denoted by $(X, S_\Delta, \nu_\Delta)$, while ν_Δ is called the *uniform Loeb measure* with weight Δ . In the case of uniform Loeb measures, every internal function $F : X \rightarrow {}^*\mathbb{R}$ is simple and \mathcal{A} -measurable; moreover, $\int_A F d\nu = \Delta \sum_{x \in A} F(x)$ for all $A \in \mathcal{A}$.

The Loeb measure ν_Δ is finite provided that $\Delta|X|$ is a limited real. If $\Delta := |X|^{-1}$ then the Loeb measure space $(X, S_\Delta, \nu_\Delta)$ is called *canonical* and denoted by (X, S, ν_L) or (X, S^X, ν_L^X) . In the case of a finite Loeb measure, if $F : X \rightarrow {}^*\mathbb{R}$ is an \mathcal{S} -integrable lifting of a function $f : X \rightarrow \overline{\mathbb{R}}$ then

$$\int_X f d\nu_\Delta = \int_X \circ \left(\Delta \sum_{x \in X} F(x) \right)$$

by Theorem 6.3.8.

6.3.11. We now abstract Theorem 6.3.8 to the case of an infinite Loeb measure ν_Δ .

Suppose that $(X, S_\Delta, \nu_\Delta)$ is a Loeb measure space, and $M \in S_\Delta$ enjoys the property: there is an increasing sequence $(M_n)_{n \in \mathbb{N}}$ of internal sets satisfying the two conditions: $M = \bigcup_{n \in \mathbb{N}} M_n$ and $\Delta|M_n| \ll +\infty$ for all $n \in \mathbb{N}$.

In this case we let S_Δ^M stand for the σ -algebra $\{A \cap M : A \in S_\Delta\}$ of subsets of M and denote the restriction of ν_Δ to S_Δ^M by ν_Δ^M . The space $\Xi := (M, S_\Delta^M, \nu_\Delta^M)$ is a σ -finite subspace of the Loeb measure space $(X, S_\Delta, \nu_\Delta)$. We also let $\mathfrak{A} := {}^*\mathcal{P}(X)$ stand for the set of all internal subsets of X .

An internal function $F : X \rightarrow {}^*\mathbb{R}$ is \mathcal{S}_M -integrable, provided that the following hold:

- (1) $\Delta \sum_{\xi \in X} |F(\xi)| \ll \infty$;
- (2) $(\forall A \in \mathfrak{A})(\Delta|A| \approx 0 \rightarrow \Delta \sum_{\xi \in A} |F(\xi)| \approx 0)$;
- (3) $(B \in \mathfrak{A} \wedge B \subset X - M) \rightarrow \Delta \sum_{\xi \in B} |F(\xi)| \approx 0$.

If $X = M$ and M is an internal set then $X = M_n$ for some $n \in \mathbb{N}$ by ω_1 -saturation. In particular, $\nu_\Delta(X) < +\infty$ and so, by 6.3.7, \mathcal{S}_M -integrability coincides in this case with \mathcal{S} -integrability.

An internal function $F : X \rightarrow {}^*\mathbb{R}$ is a *lifting* of a function $f : X \rightarrow {}^*\overline{\mathbb{R}}$ provided that $f(\xi) = {}^\circ F(\xi)$ for ν_Δ^M -almost all $\xi \in M$.

6.3.12. We introduce some notation. Let \mathcal{L} be the hyperfinite-dimensional internal vector space of functions $F : X \rightarrow {}^*\mathbb{R}$ furnished with the norm $\|F\| := \Delta \sum_{\xi \in X} |F(\xi)|$. If $A \in \mathfrak{A}$ then $\|F\|_A := \|F\chi_A\| := \Delta \sum_{\xi \in A} |F(\xi)|$, where χ_A is the *characteristic function* of A .

Recall that $\text{ltd}(\mathcal{L})$ is the external subspace of \mathcal{L} comprising the elements of limited norm, and $\mu(\mathcal{L}) \subset \text{ltd}(\mathcal{L})$ is the monad of \mathcal{L} comprising the elements of infinitesimal norm (see 6.1.1). The nonstandard hull $\mathcal{L}^\# := \text{ltd}(\mathcal{L})/\mu(\mathcal{L})$ is a nonseparable Banach space (in the case when the internal cardinality $|X|$ of X is an unlimited number). Denote by $\mathcal{S}(M)$ the subspace of $\text{ltd}(\mathcal{L})$ comprising the \mathcal{S}_M -integrable functions.

In the course of this subsection the set M is fixed, so we will write \mathcal{S} instead of $\mathcal{S}(M)$ and \mathcal{S}_M and abbreviate ν_Δ^M to ν_Δ . Finally, we let $F \sim G$ symbolize the relation $\|F - G\| \approx 0$.

- (1) *If $F \in \mathcal{L}$ is an arbitrary function $F \in \mathcal{L}$ then from $\|F\| \approx 0$ it follows that $F(\xi) \approx 0$ for ν_Δ -almost all ξ .*

\triangleleft Assume that there is some A in S_Δ such that $\nu_\Delta(A) > 0$ and ${}^\circ F(\xi) \neq 0$ for all $\xi \in A$. Show that in this case there is some internal $B \in \mathfrak{A}$ satisfying the same conditions.

Indeed, if $\nu_\Delta(A) \leq +\infty$ then $\nu_\Delta(A) = \sup\{\nu_\Delta(B) : B \in \mathfrak{A}, B \subset A\}$ by Theorem 6.3.4. If $\nu_\Delta(A) = +\infty$ then, appealing again to Theorem 6.3.4, note that either there is an internal subset B of A satisfying $\nu_\Delta(B) = +\infty$ or there is a sequence (A_n) of internal sets such that $A \subset \bigcup_{n=0}^\infty A_n$ and $\nu_\Delta(A_n) < +\infty$ for all $n \in \mathbb{N}$. In the last case $\nu_\Delta(A \cap A_n) \rightarrow +\infty$, and so there is some n satisfying $\nu_\Delta(A \cap A_n) > 0$, and again we may apply Theorem 6.3.4.

Indeed, $\nu_\Delta(B) > 0$ for some $B \in \mathfrak{A}$. Since $T := \{|F(\xi)| : \xi \in B\}$ is an internal set, it follows that $\alpha := {}^\circ(\inf T) > 0$ and ${}^\circ\|F\| \geq {}^\circ(\alpha\Delta|B|) = {}^\circ\alpha\nu_\Delta(B) > 0$. \triangleright

- (2) *If $F \in \mathcal{S}$ and $G \sim F$ then $G \in \mathcal{S}$.*

\triangleleft If A satisfies one of the conditions (2) or (3) in the definition of 6.3.11 then $\|F\|_A \approx 0$, and since $\|F - G\| \approx 0$, it follows that $\|F - G\|_A \approx 0$. Consequently, $\|G\|_A \approx 0$. \triangleright

The next result follows from similar arguments.

- (3) *The nonstandard hull $\mathcal{S}^\#$ is a closed subspace of the Banach space $\mathcal{L}^\#$.*

6.3.13. Theorem. A function $f : X \rightarrow {}^*\overline{\mathbb{R}}$ is ν_Δ^M -integrable if and only if f has an \mathcal{S}_M -integrable lifting F . In this event

$$\int_M f d\nu_\Delta^M = \circ \left(\Delta \sum_{\xi \in X} F(\xi) \right).$$

◁ It suffices clearly to prove the theorem for f a positive function. Accordingly, suppose that $f \in L_1(\nu_\Delta)$ and $f \geq 0$. Put $f_n := f\chi_{M_n}$. The sequence (f_n) increases and converges pointwise to the integrable function f , so

$$(1) \int_M f_n d\nu_\Delta \rightarrow \int_M f d\nu_\Delta.$$

This implies in the limit that

$$(2) \int_M |f_n - f_m| d\nu_\Delta \rightarrow 0, \quad n, m \rightarrow +\infty.$$

Since the support of f_n lies in the internal set M_n of finite measure, by Theorem 6.3.8 we may find an \mathcal{S} -integrable lifting F_n of f_n vanishing beyond M_n (since $F_n(\xi) = 0$ for $\xi \in X - M_n$ by 6.3.11 (3)). Moreover, we have (cf. 6.3.11)

$$(3) \int_M f_n d\nu_\Delta^M = \circ \left(\Delta \sum_{\xi \in X} F_n(\xi) \right).$$

It now follows from (2) that $\circ \|F_n - F_m\| \rightarrow 0$ as $m, n \rightarrow +\infty$. By 6.3.12 (3), there is an internal function $F \in \mathcal{S}$ satisfying $\circ \|F_n - F\| \rightarrow 0$. Clearly, $\circ \|F_n\| \rightarrow \circ \|F\|$. Passage to the limit in (3) together with (1) implies that (3) holds also for F and f .

It remains to show that $f(\xi) = \circ F(\xi)$ for ν_Δ -almost all ξ . To this end, take an arbitrary natural $k > 0$. Clearly, $\circ \|F_n\chi_{M_k} - F\chi_{M_k}\| \rightarrow 0$ as $n \rightarrow \infty$. If $n > k$ then $M_n \supset M_k$, implying that $f_n\chi_{M_k} = f\chi_{M_k}$. Consequently, the relation $F_n\chi_{M_k} \approx F_k\chi_{M_k}$ holds ν_Δ -almost everywhere. Hence, the relation $F_n\chi_{M_k} - F\chi_{M_k} \approx F_k\chi_{M_k} - F\chi_{M_k}$ also holds ν_Δ -almost everywhere.

Since the functions in the last relation are \mathcal{S} -integrable and supported in the set M_k of finite measure, it follows that $\circ \|F_n\chi_{M_k} - F\chi_{M_k}\| = \circ \|F_k\chi_{M_k} - F\chi_{M_k}\|$. Passing here to the limit as $n \rightarrow \infty$, infer that $\circ \|F_k\chi_{M_k} - F\chi_{M_k}\| = 0$ for any k . By 6.3.12(1), conclude now that the relation $F_k\chi_{M_k} \approx F\chi_{M_k}$ holds ν_Δ -almost everywhere.

Consequently, the relation $f\chi_{M_k} \approx F\chi_{M_k}$ holds ν_Δ -almost everywhere, implying that the relation $f \approx F|_M$ also holds ν_Δ -almost everywhere because $M = \bigcup_{k=0}^{\infty} M_k$.

Suppose now that $F \in \mathcal{S}$ and $f : M \rightarrow \mathbb{R}$ is such that $f(\xi) = \circ F(\xi)$ for almost all $\xi \in M$. Show that $f \in L_1(\nu_\Delta)$ and (3) holds for F and f . Put $F_n := F\chi_{M_k}$. Then the relation $f_n := f\chi_{M_k} \approx F_n$ holds ν_Δ -almost everywhere. Hence, we may

apply Theorem 6.3.8 since $\nu_\Delta(M_n)$ is finite. This yields

$$\int_M |f_n| d\nu_\Delta = \circ\|F_n\| \leq \circ\|F\|.$$

Since $(|f_n|)$ increases and converges to $|f|$, it follows from the Monotone Convergence Theorem that $f \in L_1(\nu_\Delta)$. By the above, there is an \mathcal{S} -integrable internal function $G : X \rightarrow {}^*\mathbb{R}$ satisfying $\int_M |f_n| d\nu_\Delta = \circ\|G\|$.

To prove (3) for F and f , it remains to show that $\|F - G\| \approx 0$. Note first that if $A \in \mathfrak{A}$ and $A \subset M$ then $A \subset M_n$ for some number n by ω_1 -saturation. Take such a set A and such a number n . Put $F_M := F\chi_M$. Then the relations $f \approx G_M$ and $f \approx F_M$ hold ν_Δ -almost everywhere and so the relation $F_{M_n} \approx G_{M_n}$ holds ν_Δ -almost everywhere too. But then $\|F - G\|_A \approx 0$ because F and G are \mathcal{S} -integrable.

Consider the family of formulas $\Gamma_{m,n}(A) := \{A \in \mathfrak{A} \wedge M_n \subset A \wedge \|F - G\|_A \leq m^{-1}\}$. To each natural $N \in \mathbb{N}$, there is some $A \in \mathfrak{A}$ such that $\Gamma_{m,n}(A)$ holds for all $n, m \leq N$. Hence, there is some $A \in \mathfrak{A}$ satisfying all formulas $\Gamma_{m,n}(A)$. But then $M \subset A$ and $\|F - G\|_A \approx 0$. Since $F, G \in \mathcal{S}$; therefore, $\|F - G\|_{X-A} \approx 0$. Consequently, $\|F - G\| \approx 0$. \triangleright

6.3.14. Take $p \in [1, +\infty)$, and let \mathcal{L}_p stand for the internal vector space of functions $F : X \rightarrow {}^*\mathbb{R}$ under the norm

$$\|F\|_p = \left(\Delta \sum_{\xi \in X} |F(\xi)|^p \right)^{1/p}.$$

Sometimes this space and its norm are denoted lavishly by $\mathcal{L}_{p,\Delta}^X$, and $\|\cdot\|_{p,\Delta}$. Given $F, G \in \mathcal{L}_p$, we write $F \underset{p}{\sim} G$ whenever $\|F - G\|_p \approx 0$. The nonstandard hull $\mathcal{L}_p^\#$ is defined in exactly the same way as in Section 6.1.

If $A \in \mathfrak{A}$ then $\|F\|_{p,A} = \|F\chi_A\|_p$. Denote by $\mathcal{S}_p(M)$ the subspace of \mathcal{L}_p consisting of the functions $F \in \mathcal{L}_p$ such that the power $|F|^p$ is \mathcal{S}_M -integrable.

We will write simply \mathcal{S}_p , omitting M when this leads to no confusion. Since $\|F\|_{p,A} = \| |F|^p \|_A$ for every internal function F and every $A \in \mathfrak{A}$; therefore, Propositions 6.3.12 (1)–(3) remain valid on replacing \mathcal{L} , \mathcal{S} , and $\|\cdot\|$ with \mathcal{L}_p , \mathcal{S}_p , and $\|\cdot\|_p$ respectively.

We also define the complex spaces \mathcal{L}_p and \mathcal{S}_p in a completely analogous manner. Furthermore, if $F : X \rightarrow \mathbb{C}$ is an internal function then $F = \text{Re } F + i \text{Im } F$ and

$$\| \text{Re } F \|_{p,A}, \| \text{Im } F \|_{p,A} \leq \| F \|_{p,A} \leq \| \text{Re } F \|_{p,A} + \| \text{Im } F \|_{p,A}$$

for every $A \in \mathfrak{A}$.

From the last inequalities it follows that $F \in \mathcal{L}_p(\mathcal{S}_p)$ if and only if $\operatorname{Re} F \in \mathcal{L}_p(\mathcal{S}_p)$ and $\operatorname{Im} F \in \mathcal{L}_p(\mathcal{S}_p)$.

If $f : M \rightarrow \mathbb{F}$, with \mathbb{F} a basic field of scalars (i.e., \mathbb{R} or \mathbb{C}) then $f \in L_p(\Xi)$ if and only if there is a lifting $F : X \rightarrow {}^*\mathbb{F}$ of f such that $F \in \mathcal{S}_p(M)$. Furthermore, $\|f\|_p = {}^\circ\|F\|_p$.

6.3.15. COMMENTS.

(1) The construction of Loeb measure was implemented in [322]. The content of 6.3.1–6.3.10 is well known; see [3, 69]. Theorem 6.3.4 in the case of finite measure belongs to Loeb [322]. Henson [166] established uniqueness in the case of infinite measure as well as the property 6.3.4 (4); the results of 6.3.2 and 6.3.3 are extracted from [166].

(2) Theorem 6.3.6 belongs to Loeb [322, 323]. A similar characterization is available for measurable mappings with values in a complete separable metric space (see [12, 322]). The definition of \mathcal{S} -integrable function (6.3.7) was introduced by Loeb [323]; somewhat earlier Anderson considered the equivalent condition 6.3.7 (1) in [11].

(3) The concept of a σ -finite subspace of a Loeb measure space was introduced by Gordon in [145]. The same article reveals Theorem 6.3.13. In 6.3.11–6.3.14, we proceed along the lines of [146].

6.4. Hyperapproximation of Measure Space

The aim of this section is to show that each standard σ -finite measure space embeds in the Loeb measure space of an appropriate hyperfinite uniform measure space. Some arguments below imply that the nonstandard universe enjoys Nelson's idealization principle.

6.4.1. We now prove that, given a σ -finite measure space $(\mathfrak{X}, \Omega, \mu)$, we may construct a Loeb measure space $(X, S_\Delta, \nu_\Delta)$ and a σ -finite subspace $(M, S_\Delta^M, \nu_\Delta^M)$ of $(\mathfrak{X}, \Omega, \mu)$ so that $X \subset {}^*\mathfrak{X}$ and to each $p \in [1, \infty)$ there corresponds an isometric embedding $J_p : L_p(\mu) \rightarrow L_p(\nu_\Delta^M)$. Furthermore, if $f \in L_p(\mu)$ then the internal function $F := {}^*f|_X$ belongs to $\mathcal{S}_p(M)$ and serves as a lifting of $J_p(f)$. This implies in particular that

$$\int_{\mathfrak{X}} f d\mu = {}^\circ \left(\Delta \sum_{\xi \in X} {}^*f(\xi) \right)$$

for $f \in L_p(\mu)$ (more precisely, for each member of the coset f).

Let (Y, Σ, μ) be a standard measure space. An element ξ in *Y is *random* provided that ξ belongs to no μ -negligible set. In other words, an element $\xi \in {}^*Y$ is random whenever $\xi \notin {}^*A$ for each standard $A \in \Sigma$ satisfying $\mu(A) = 0$.

(1) *Almost all elements of *Y are random. More precisely, there is an internal set $B \in \Sigma$ such that $\mu({}^*Y - B) = 0$ and every member of B is random.*

◁ Let J be the ideal of μ -negligible sets. By idealization, there is a hyperfinite set $\mathcal{M} \subset {}^*J$ such that ${}^*A \in \mathcal{M}$ for all standard $A \in J$. Put $X := \bigcup \mathcal{M}$. Then $X \in {}^*Y$, $\mu(X) = 0$. Obviously, if $\xi \in {}^*Y - X$ then ξ is a random element and $Y - X$ is a set of full measure. ▷

(2) *Assume that the available nonstandard universe enjoys the concurrence principle. Let (X, \mathcal{Y}, μ) be a standard measure space. Then there is some $\xi \in {}^*X$ satisfying*

$$(\forall Y \in \mathcal{Y})(\mu(Y) = 0 \rightarrow \xi \in {}^*Y).$$

◁ Consider the internal class $\mathfrak{N} := \{Y \in \mathcal{Y} : \mu(Y) = 0\}$. It is possible to show by concurrence that there is a hyperfinite family $G := \{G_n : n < \lambda\}$, with $\lambda \in {}^*\mathbb{N}$, such that $G \subset \mathfrak{N}$ and ${}^*A \in \mathfrak{N} \leftrightarrow {}^*A \in G$. By transfer, $Y_0 := \bigcup \{G_n : n < \lambda\} \in {}^*\mathcal{Y}$. Hence, $\mu(Y_0) = 0$, implying that $\mu({}^*X - Y_0) = \mu({}^*X) = \mu(X)$. Therefore, ${}^*X - Y_0 \neq \emptyset$ (we presume of course that $\mu(X) > 0$). Clearly, each member ${}^*X - Y_0$ satisfies the claim. ▷

The concept of random element extends to τ -standard measure spaces. Suppose to this end that τ is an admissible element and (Y, Σ, λ) is a τ -standard σ -finite probability space (i.e., $\lambda(Y) = 1$). An element $y \in Y$ is τ -random provided that $y \notin A$ for every τ -standard λ -negligible set $A \in \Sigma$.

In particular, if τ is a standard real implying that (Y, Σ, λ) is a standard probability space then every τ -random element $y \in {}^*Y$ is random. Of course, τ -random elements are defined also in each standard space (Y, Σ, λ) given a nonstandard τ . Claim (1) remains true in the case of τ -standard space.

(3) *There is an internal set $B \in \Sigma$ such that $\mu({}^*Y - B) = 0$ and the elements of B are all τ -random.*

◁ The proof repeats the arguments of (1); however, instead of the idealization principle for τ -standard sets we appeal to the relative idealization principle. ▷

6.4.2. Suppose now that (Y, Σ, λ) is the product of the τ -standard probability spaces $(Y_1, \Sigma_1, \lambda_1)$ and $(Y_2, \Sigma_2, \lambda_2)$. If $y = (y_1, y_2)$ is a τ -random element of Y then it is easy to see that y_i is a τ -standard element of Y_i . The converse fails. For example, if $y_1 = y_2$ and the measure of the diagonal in $I_Y := \{(y, y) : y \in Y\}$ is zero then (y_1, y_2) is not τ -random even if y_1 is τ -random, since (y_1, y_2) belongs to I_Y , a τ -standard negligible set.

(1) *If y_1 is a τ -random element of Y_1 and y_2 is a (τ, y_1) -random element of Y_2 then (y_1, y_2) is a τ -random element of $Y = Y_1 \times Y_2$.*

◁ Let $A \subset Y_1 \times Y_2$ be a τ -standard set with $\lambda(A) = 0$. Then $A_{z_1} := \{z_2 \in Y_2 : (z_1, z_2) \in A\}$ is (τ, z_1) -standard for all $z_1 \in Y_1$. Put $C := \{z_1 \in Y_1 : \lambda_2(A_{z_1}) = 0\}$.

Then C is τ -standard and $\lambda_1(C) = 1$ by the Fubini theorem. By hypothesis y_1 is a τ -random element and so $y_1 \in C$. Hence, $\lambda_2(A_{y_1}) = 0$ and $y_2 \notin A_{y_1}$ because A_{y_1} is a (τ, y_1) -standard set. Thus, $(y_1, y_2) \notin A$, implying that (y_1, y_2) is a τ -random element. \triangleright

(2) Let (Y, Σ, λ) be a τ -standard probability space. A countable family $(y_n)_{n \in \mathbb{N}}$ is an *independent sequence* of τ -random elements of Y if (y_n) is a τ -random element of the τ -standard space $(Y^{*\mathbb{Z}}, \lambda^\infty)$, with λ^∞ the countable power of λ .

6.4.3. Take an independent sequence $(y_n)_{n \in \mathbb{N}}$ of τ -random elements of Y . We are interested as to whether to following representation holds:

$$(1) \int_Y f d\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(y_k)$$

with f an arbitrary τ -standard function.

(2) Let (Y, Σ, λ) be a probability space, and let λ^∞ stand for the countable power of λ on $Y^{\mathbb{Z}}$. Given an integrable function $f : Y \rightarrow \mathbb{R}$, denote by \mathcal{A}_f the set of sequences $(y_n)_{n \in \mathbb{Z}} \in Y^{\mathbb{Z}}$ satisfying (1). Then $\lambda^\infty(\mathcal{A}_f) = 1$.

\triangleleft Let $T : Y^{\mathbb{Z}} \rightarrow Y^{\mathbb{Z}}$ be the *shift operator* or, *Bernoulli automorphism*:

$$T((y_n)_{n \in \mathbb{Z}}) := (y'_n)_{n \in \mathbb{Z}}, \quad y'_n := y_{n+1}.$$

It is well known that T is an *ergodic* operator (cf. [245, Chapter 8, § 1, Theorem 1]), i.e.,

$$\int_{Y^{\mathbb{Z}}} \varphi d\lambda^\infty = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k \bar{y})$$

for each function $\varphi \in L_1(\lambda^\infty)$ and almost all $\bar{y} \in Y^{\mathbb{Z}}$.

Let the mapping $\Pi_0 : Y^{\mathbb{Z}} \rightarrow Y$ be defined by the formula $\Pi_0((y_n)_{n \in \mathbb{Z}}) := y_0$. By the definition of λ^∞ , the mapping Π_0 is measure-preserving. Hence, if $\varphi = f \circ \Pi_0$ then $\int_{Y^{\mathbb{Z}}} \varphi d\lambda^\infty = \int_Y f d\lambda$. Moreover, $\varphi(T^k \bar{y}) = (f \circ \Pi_0)(T^k \bar{y}) = f(y_k)$, yielding the claim. \triangleright

(3) **Theorem.** If $(y_n)_{n \in \mathbb{N}}$ is an independent sequence of τ -random elements of Y then (1) holds for every τ -standard function $f : Y \rightarrow {}^*\mathbb{R}$ belonging to $L_1(\lambda)$.

\triangleleft By relative transfer, from (2) we infer that \mathcal{A}_f is a τ -standard set of full measure. Hence, every independent sequence $(y_n)_{n \in \mathbb{N}}$ of τ -standard elements lies in \mathcal{A}_f . This implies the validity of (1). \triangleright

6.4.4. Theorem. If (Y, Σ, δ) is a τ -standard finite measure space then there is an internal hyperfinite set $Y_0 \subset Y$ such that

$$\int_Y f d\delta \stackrel{\tau}{\approx} \frac{\delta(Y)}{|Y_0|} \sum_{y \in Y_0} f(y)$$

for every τ -standard integrable function $f : Y \rightarrow {}^*\mathbb{R}$.

◁ Note first that $\delta(Y)$ is a τ -standard, in general, unlimited hyperreal, since the condition that δ is finite means exactly that $\delta(Y) \neq \infty$. Pass from the measure space (Y, Σ, δ) to the probability space (Y, Σ, λ) by setting $\lambda := \frac{1}{\delta(Y)}\delta$. Obviously, the integrals with respect to δ and λ are related as follows:

$$\int_Y f d\delta = \delta(Y) \int_Y f d\lambda.$$

Let $\bar{y} := (y_n)_{n \in {}^*\mathbb{Z}}$ be an independent sequence of τ -random elements of (Y, Σ, λ) and let N be some (τ, \bar{y}) -infinite hypernatural. Since the sequence on the right side of 6.4.3(1) is (τ, \bar{y}) -standard; therefore, from 4.6.4(1) it follows that

$$\int_Y f d\lambda \stackrel{(\tau, \bar{y})}{\approx} \frac{1}{N} \sum_{k=0}^{N-1} f(y_k).$$

Put $Y_0 := \{y_0, \dots, y_{N-1}\}$. It remains to use the above-mentioned relation between the integrals with respect to δ and λ together with the fact that $\alpha \stackrel{(\tau, \bar{y})}{\approx} \beta$ yields $\alpha \stackrel{\tau}{\approx} \beta$ (see 4.6.2). ▷

6.4.5. We return to considering a standard σ -finite measure space $(\mathfrak{X}, \Omega, \mu)$. Thus, there is an increasing sequence of sets $\mathfrak{X}_n \in \Omega$ such that $\mu(\mathfrak{X}_n) < +\infty$ for all $n \in \mathbb{N}$ and $\mathfrak{X} = \bigcup_{n \in \mathbb{N}} \mathfrak{X}_n$. If Ω_n stands for the σ -algebra $\{A \cap \mathfrak{X}_n : A \in \Omega\}$ and μ_n is the restriction of μ to Ω_n then

$$\int_{\mathfrak{X}} f d\mu = \lim_{n \rightarrow \infty} \int_{\mathfrak{X}_n} f_n d\mu_n$$

for every integrable function $f : \mathfrak{X} \rightarrow \mathbb{R}$, with $f_n := f|_{\mathfrak{X}_n}$.

Theorem. *There exist an internal hyperfinite set $X \subset {}^*\mathfrak{X}$ and a hyperreal $\Delta \in {}^*\mathbb{R}$ such that*

$$\int_{\mathfrak{X}} f d\mu = \overset{\circ}{\left(\Delta \sum_{\xi \in X} {}^*f(\xi) \right)}$$

for every standard function $f \in L_1(\mu)$.

◁ Let τ be an unlimited hypernatural. Put $(Y, \Sigma, \delta) := ({}^*\mathfrak{X}_\tau, {}^*\Omega_\tau, {}^*\mu_\tau)$. Then (Y, Σ, δ) satisfies the hypotheses of Theorem 6.4.4. Clearly, ${}^*f_\tau$ is a τ -standard integrable function on ${}^*\mathfrak{X}_\tau$. Since $Y_0 \subset {}^*\mathfrak{X}_\tau$ (see the definition of Y_0 in 6.4.4), it

follows that $*f_\tau|_{Y_0} = *f|_{Y_0}$. Note also that from the above limit relation it follows that

$$\int_{\mathfrak{X}} f d\mu = \int_{*\mathfrak{X}_\tau} {}^{\circ} f_\tau d\mu_\tau.$$

This together with Theorem 6.4.4 yields the claim on putting $\Delta := \delta(Y)|Y_0|^{-1}$ and $X := Y_0$ and using the fact that $\int_{\mathfrak{X}} f d\mu$ is a standard real. \triangleright

6.4.6. In the sequel we will need some facts from the theory of normed Boolean algebras. All these facts can be found in [505]. Let \mathcal{B} be a Boolean algebra, and let m be a strictly positive finitely additive measure on \mathcal{B} . Then \mathcal{B} satisfies the *countable chain condition* or, in the literature of Russian provenance, \mathcal{B} has *countable type*; i.e., every disjoint subset E of \mathcal{B} is at most countable. If \mathcal{B} satisfies the countable chain condition then \mathcal{B} is complete. A complete Boolean algebra with a strictly positive measure is said to be *normed*. Recall that a *measure* is by definition a positive countably additive function on \mathcal{B} .

Let (\mathcal{B}_1, m_1) and (\mathcal{B}_2, m_2) be normed Boolean algebras. Every measure-preserving homomorphism $\varphi : \mathcal{B}_1 \rightarrow \mathcal{B}_2$ is *completely additive*; i.e., $\varphi(\sup E) = \sup \varphi(E)$ for all $E \subset \mathcal{B}_1$. This implies that $\varphi(\mathcal{B}_1)$ is a regular subalgebra of \mathcal{B}_2 , i.e., for every set A in $\varphi(\mathcal{B}_1)$ the greatest lower bound of A and the least upper bound of A , if exist in \mathcal{B}_2 , belong to $\varphi(\mathcal{B}_1)$.

Suppose now that (Y, Σ, δ) is a finite measure space, and put $n(\Sigma) := \{c \in \Sigma : \delta(c) = 0\}$. Then $\bar{\Sigma} := \Sigma/n(\Sigma)$ is a normed Boolean algebra with strictly positive measure $\bar{\delta}$ such that $\bar{\delta}([c]) = \delta(c)$, with $[c]$ the coset of an element $c \in \Sigma$ in $\bar{\Sigma}$. The normed algebra $\bar{\Sigma}$ is the *Lebesgue algebra* of (Y, Σ, δ) in common parlance.

With each measurable function $f : Y \rightarrow \mathbb{R}$ we associate an increasing right-continuous family $(e_f^t)_{t \in \mathbb{R}}$ of elements in $\bar{\Sigma}$ which is called the *characteristic* or *resolution of the identity* for f and which is defined as $e_f^t := [\{y : f(y) \leq t\}]$. Note that $\sup(e_f^t) = \mathbf{1}_{\bar{\Sigma}}$ and $\inf(e_f^t) = \mathbf{0}_{\bar{\Sigma}}$.

(1) A measurable function f is integrable if and only if the integral $\int_{-\infty}^{\infty} t d\bar{\delta}(e_f^t)$ converges and

$$\int_Y f d\delta = \int_{-\infty}^{\infty} t d\bar{\delta}(e_f^t).$$

\triangleleft This simple fact is well known; cf. [505, Chapter VI, § 3]. \triangleright

(2) Assume that (Y, Σ, δ) is a standard finite measure space, $Y_0 \subset *Y$ satisfies the hypotheses of Theorem 6.4.4 (with τ standard), $\lambda := \delta(Y)|Y_0|^{-1}$ and $(Y_0, S_\lambda, \nu_\lambda)$ is the corresponding Loeb measure space. Then the mapping

$\psi : \bar{\Sigma} \rightarrow \bar{S}_\lambda$, acting by the rule $\psi([c]) := [{}^*c \cap Y_n]$ ($c \in \Sigma$), is a measure-preserving monomorphism.

◁ This is immediate from Theorem 6.4.4 on applying the formula of this theorem to the characteristic functions of members of Σ . ▷

(3) In the context of (2) suppose additionally that $h \in L_1(\delta)$ and $H := {}^*h|_{Y_0}$. Suppose also that the function $\tilde{h} : Y_0 \rightarrow \bar{\mathbb{R}}$ satisfies the equality $\tilde{h}(y) = {}^\circ h(y)$ for all $y \in Y_0$. Then $\tilde{h} \in L_1(\nu_\lambda)$, $\int_Y h d\delta = \int_{Y_0} \tilde{h} d\nu_\lambda$, and H is an \mathcal{S} -integrable function.

◁ Given a standard $t \in \mathbb{R}$, put $C_t := \{y \in Y : h(y) \leq t\}$, $c_t := [C_t] \in \bar{\Sigma}$, $E_t := \{y \in Y_0 : \tilde{h}(y) \leq t\}$, and $e_t := [E_t] \in \bar{S}_\lambda$. Then $(c_t)_{t \in \mathbb{R}}$ is the resolution of the identity for h in $\bar{\Sigma}$ and $(e_t)_{t \in \mathbb{R}}$ is the resolution of the identity for \tilde{h} .

Note that the mapping $\psi : \bar{\Sigma} \rightarrow \bar{S}_\lambda$, defined in (2), is a full monomorphism (that is, it preserves suprema and infima of arbitrary sets). So, if $\tilde{e}_t := \psi(c_t)$ then $(\tilde{e}_t)_{t \in \mathbb{R}}$ is a resolution of the identity.

By transfer, $\tilde{e}_t = \{y \in Y_0 : {}^*h(y) \leq t\}$. Using the definition of \tilde{h} , infer that $e_{t_1} < \tilde{e}_{t_2}$ and $\tilde{e}_{t_1} < e_{t_2}$ for all standard $t_1 < t_2$. This, together with the right-continuity of the families $(e_t)_{t \in \mathbb{R}}$ and $(\tilde{e}_t)_{t \in \mathbb{R}}$, implies that $e_t = \tilde{e}_t$ for all t . We now obtain the first two claims from (1) and the fact that ψ is measure-preserving. The third claim follows from 6.3.7 (1). ▷

6.4.7. From 6.4.1 it is clear that

$$\mu(A) = {}^\circ(\Delta|X \cap {}^*A)$$

for all standard $A \in \Omega$ (here, as usual, ${}^\circ t := +\infty$ whenever $t \in {}^*\mathbb{R}$ and $t \approx +\infty$.) This, together with the inequality $\mu({}^*\mathfrak{X}_n) < +\infty$, implies that the triple $\Xi = (M, S_\Delta^M, \nu_\Delta^M)$, with $M_n := X \cap {}^*\mathfrak{X}_n$ for $n \in \mathbb{N}$ and $M := \bigcup_{n \in \mathbb{N}} M_n$, is a σ -finite subspace of the Loeb measure space $(X, S_\Delta, \nu_\Delta)$.

Theorem. Let $(\mathfrak{X}, \Omega, \mu)$ be a standard σ -finite measure space. Assume also that $\mathfrak{X} = \bigcup_{n \in \mathbb{N}} \mathfrak{X}_n$ and $\mu(\mathfrak{X}_n) < +\infty$ for all $n \in \mathbb{N}$. Assume further that $X \subset {}^*\mathfrak{X}$ and $\Delta \in {}^*\mathbb{R}$ satisfy the conditions of Theorem 6.4.5. Put $M_n := X \cap {}^*\mathfrak{X}_n$ and $M := \bigcup_{n \in \mathbb{N}} M_n$. If for all $p \in [1, +\infty)$ and $f \in L_p(\mu)$ the internal function $F(f) := {}^*f|_X$ belongs to $\mathcal{S}_p(M)$ and $j_p(f) = {}^\circ F(f)$, then $j_p : L_p(\mu) \rightarrow L_p(\nu_\Delta^M)$ is an isometric embedding. In particular, if $f \in L_1(\mu)$ then

$$\int_{\mathfrak{X}} f d\mu = \int_M j_1(f) d\nu_\Delta^M.$$

◁ It suffices to prove the theorem for $p = 1$ and $f \geq 0$.

We consider the finite Loeb measure space $(M_n, S_{\Delta_n}, \nu_{\Delta_n})$, on assuming $\Delta_n = \mu(\mathfrak{X}_n)|M_n|^{-1} \approx \Delta|M_n||M_n|^{-1} = \Delta$. In this event the conditions of 6.4.6 (3) hold on replacing (Y, Σ, δ) with $(\mathfrak{X}_n, \Omega_n, \mu_n)$ and Y_0 with M_n .

Put $f_n := f\chi_{\mathfrak{X}_n} : \mathfrak{X} \rightarrow \overline{\mathbb{R}}$ and $\bar{f}_n := f|_{\mathfrak{X}_n}$. By 6.4.6 (3), $*f|_{M_n}$ is an \mathcal{S} -integrable lifting of ${}^\circ(*f|_{M_n})$ and

$$\begin{aligned} \int_{\mathfrak{X}_n} \bar{f}_n d\mu_n &= \int_{M_n} {}^\circ(*f|_{M_n}) d\nu_{\Delta_n} \\ &= {}^\circ\left(\Delta_n \sum_{\xi \in M_n} * \bar{f}_n(\xi)\right) = {}^\circ\left(\Delta_n \sum_{\xi \in X} * f_n(\xi)\right). \end{aligned}$$

The last equality holds because $\frac{\Delta_n}{\Delta} \approx 1$ and $*f_n(\xi) = 0$ for $\xi \in X - M_n$. This implies also that $*f_n|_X$ is an \mathcal{S}_M -integrable lifting of $j_1(f_n)$, and $\int_{\mathfrak{X}} f_n d\mu = \int_M j_1(f_n) d\nu_{\Delta}^M$. Since $(j_1(f_n))$ increases and converges pointwise to $j_1(f)$, passage to the limit in the last equality enables us to conclude that this equality remains valid for f and $j_1(f) \in L_1(\nu_{\Delta})$. By Theorem 6.4.5,

$${}^\circ\left(\Delta \sum_{\xi \in X} |*f_n(\xi) - *f(\xi)|\right) = \int_{\mathfrak{X}} |f_n - f| d\mu \rightarrow 0.$$

Since $\mathcal{S}(M)^\#$ is closed in $\mathcal{L}^\#$ (see 6.3.12 (3)), we now conclude that $*f|_X$ belongs to $\mathcal{S}(M)$, swerving an \mathcal{S}_M -integrable lifting of the function $j_1(f)$. \triangleright

6.4.8. For many concrete σ -finite measure spaces (X, Ω, μ) there are embeddings $L_p(\mu)$ and $L_p(\nu_{\Delta}^M)$ other than we have described in the preceding subsection on using a σ -finite subspace $(M, S_{\Delta}^M, \nu_{\Delta}^M)$ of a suitable Loeb measure space $(X, S_{\Delta}, \nu_{\Delta})$. Most of these embeddings rest on constructing a measure-preserving mapping $\varphi : M \rightarrow \mathfrak{X}$. Such a mapping φ induces the embedding $j_p^\varphi : L_p(\mu) \rightarrow L_p(\nu_{\Delta}^M)$, given $p \in [1, \infty)$, by the rule $j_p^\varphi(f) := f \circ \varphi$ for all $f \in L_p(\mu)$. By 6.3.13 and 6.3.14, $j_p^\varphi(f)$ has a lifting $F \in \mathcal{S}_p(M)$ which we call the *lifting* of f . In particular, the following is valid.

(1) If $f \in L_1(\mu)$ and F is an \mathcal{S}_M -integrable lifting of f then

$$\int_{\mathfrak{X}} f d\mu = {}^\circ\left(\Delta \sum_{\xi \in X} F(\xi)\right).$$

The interesting problem is worth noting of constructing F from f . This problem was solved in an especially simple way in the preceding subsections: there is a suitable hyperfinite set $X \subset {}^*\mathfrak{X}$ satisfying

(2) $F = {}^*f|_X$ for all $f \in L_1(\mu)$.

In the general case, (2) fails even on assuming that $X \subset {}^*\mathfrak{X}$. We now consider a type of embedding for which (2) holds for a sufficiently broad class of integrable functions. However, we start with exhibiting a well-known example.

(3) Put $\mathfrak{X} := [0, 1]$, and let μ be Lebesgue measure on \mathfrak{X} . Distinguishing an arbitrary hyperreal $\Delta \approx 0$, put $N := [\Delta^{-1}]$ and $X := \{k\Delta : k = 1, \dots, N\}$. In this event, the Loeb measure space $(X, S_\Delta, \nu_\Delta)$ is in fact a probability space: $\nu_\Delta(X) = 1$, yielding $M = X$. As a mapping $\varphi : X \rightarrow \mathfrak{X}$ we take st (recall that $\text{st}(k\Delta) = {}^\circ(k\Delta)$).

It is possible to show that a subset \mathcal{A} of $[0, 1]$ is Lebesgue-measurable if and only if $\text{st}^{-1}(\mathcal{A})$ is Loeb-measurable and $\mu(\mathcal{A}) = \nu_\Delta(\text{st}^{-1}(\mathcal{A}))$. The equality (2) fails to hold for each Lebesgue-integrable function f . It is easy to see on assuming that $\Delta \in {}^*\mathbb{Q}$ and considering the Dirichlet function.

However, if f is Riemann-integrable on $[0, 1]$ and F satisfies (2), then (1) holds by 2.3.16. We will show that in this case $F := {}^*f|_X$ really is an \mathcal{S} -integrable lifting of f . Since f is bounded and the internal uniform measure ν_Δ is finite, it follows that F satisfies 6.3.7 (1). Hence, F is \mathcal{S} -integrable. If \mathcal{A} is the set of discontinuity points of f then $\mu(\mathcal{A}) = 0$ because f is Riemann-integrable. If $k\Delta \in X - \text{st}^{-1}(\mathcal{A})$ then f is continuous at the point ${}^\circ(k\Delta)$, so that ${}^*f(k\Delta) \approx f({}^\circ(k\Delta))$. Thus, ${}^\circ F(\xi) = f(\text{st}(\xi))$ for almost all $\xi \in X$. Consequently, F is a lifting of $f \circ \text{st}$, implying that F is a lifting of f .

6.4.9. Suppose now that \mathfrak{X} is a separable locally compact Hausdorff topological space, μ is a Borel measure on \mathfrak{X} finite at compact sets (μ is regular because \mathfrak{X} is separable), and Ω is the completion of the σ -algebra of Borel sets with respect to μ . Suppose further that $\mathfrak{X} = \bigcup_{n \in \mathbb{N}} \mathfrak{X}_n$, where \mathfrak{X}_n is a compact and $\mu(\mathfrak{X}_n) < +\infty$ for all $n \in \mathbb{N}$. Then $\text{nst}({}^*\mathfrak{X}) = \bigcup {}^*\mathfrak{X}_n$. Recall that the mapping $\text{st} : \text{nst}({}^*\mathfrak{X}) \rightarrow \mathfrak{X}$ is determined by the condition $\text{st}(x) \approx x$ for all $x \in \text{nst}({}^*\mathfrak{X})$ (see 4.3.4 and 4.3.6).

(1) Suppose that X is a hyperfinite set, $j : X \rightarrow {}^*\mathfrak{X}$ is an internal mapping, $\Delta \in {}^*\mathbb{R}$, and $M := j^{-1}(\text{nst}({}^*\mathfrak{X}))$.

The triple (X, j, Δ) is a *hyperapproximant* or *hyperfinite realization* of $(\mathfrak{X}, \Omega, \mu)$ provided that the mapping $\varphi : (M, S_\Delta^M, \nu_\Delta^M) \rightarrow (\mathfrak{X}, \Omega, \mu)$ acting by the rule $\varphi := \text{st} \circ j|_M$ is measurable and measure-preserving.

Note that $M = \bigcup_{n \in \mathbb{N}} j^{-1}({}^*\mathfrak{X}_n)$, so that in this event $(M, S_\Delta^M, \nu_\Delta^M)$ is a σ -finite subspace of the Loeb measure space $(X, S_\Delta, \nu_\Delta)$.

We will formulate below one sufficient condition for the function $F := {}^*f \circ j$ be an \mathcal{S}_M -integrable lifting of f . To this end, we give the following condition expressing that f has a rather rapid decay at infinity:

$$(2) (\forall B \in {}^*\mathcal{P}(X)) \left(B \subset X - M \rightarrow \Delta \sum_{x \in B} |{}^*f(j(x))| \approx 0 \right).$$

6.4.10. Let (X, j, Δ) be a hyperapproximant to $(\mathfrak{X}, \Omega, \mu)$. Assume that $f :$

$\mathfrak{X} \rightarrow \mathbb{R}$ is a μ -integrable bounded function continuous μ -almost everywhere and satisfying the condition 6.4.9 (2). Then the function $F := {}^*f \circ j$ is an \mathcal{S}_M -integrable lifting of f and so

$$\int_{\mathfrak{X}} f \, d\mu = \overset{\circ}{\left(\Delta \sum_{x \in X} {}^*f(j(x)) \right)}.$$

\triangleleft Show first that F is \mathcal{S}_M -integrable. Note that 6.3.11 (2) holds because f is bounded, and 6.3.11 (3) follows from 6.4.9 (2). To verify 6.3.11 (1) put $M_n := j^{-1}({}^*\mathfrak{X}_n)$ and observe that ${}^*f \circ j|_{M_n}$ is an \mathcal{S} -integrable function. This is so since f is bounded and $\nu_\Delta(M_n)$ is finite (see 6.3.7 (1)). Arguing as at the end of 6.4.8 (3), conclude that ${}^*f \circ j|_{M_n}$ is a lifting of the function $f|_{\mathfrak{X}_n}$. By 6.4.6 (2) this yields

$$\int_{\mathfrak{X}_n} |f| \, d\mu = \overset{\circ}{\left(\Delta \sum_{x \in M_n} |{}^*f \circ j|(x) \right)} \leq \int_{\mathfrak{X}} |f| \, d\mu.$$

Hence, there is a standard constant C such that

$$\Delta \sum_{x \in \mathcal{D}} |F|(x) < C$$

for every internal $\mathcal{D} \subset \bigcup_{n \in \mathbb{N}} M_n$. By countable saturation, we may find an internal subset \mathcal{D} of M satisfying the last inequality. Now 6.3.11 (1) follows on applying 6.4.9 (2) to $B := X - \mathcal{D}$. Since ${}^*f \circ j|_{M_n}$ is a lifting of $f|_{\mathfrak{X}_n}$ for every $n \in \mathbb{N}$; therefore, ${}^*f \circ j$ is a lifting of f . \triangleright

The analogous assertion is, of course, valid for a bounded μ -almost everywhere continuous function $f \in L_p(\mu)$, with $p \in [1, \infty)$.

6.4.11. We now consider an example.

Let Ω be the σ -algebra of Lebesgue-measurable sets on $\mathfrak{X} := \mathbb{R}$, and let μ stand for Lebesgue measure on \mathbb{R} . Choose an infinite hypernatural $N \in {}^*\mathbb{N} - \mathbb{N}$ and an unlimited hyperreal $\Delta \in {}^*\mathbb{R} - \mathbb{R}$ such that $\Delta \approx 0$ and $N\Delta \approx +\infty$. For convenience of notation we assume that $N = 2L + 1$ and consider the hyperfinite set $X := \{k\Delta : k = -L, \dots, L\}$. Put $\mathfrak{X}_n := [-n, n]$ and consider the identity embedding $j : X \rightarrow {}^*\mathbb{R}$. Then $M_n = X \cap {}^*[-n, n]$ and $M = X \cap \bigcup_{n \in \mathbb{N}} {}^*[-n, n]$.

(1) It is easy to see that we may rewrite 6.4.9 (2) as

$$(\forall k, l) \left(|k| < |l| < L \wedge |k|\Delta \approx +\infty \rightarrow \Delta \sum_{n=k}^l |{}^*f(n\Delta)| \approx 0 \right).$$

(2) Since the last relation holds for L satisfying $L\Delta \approx +\infty$, it amounts to the equality

$$\lim_{\substack{\Delta \rightarrow 0 \\ A \rightarrow \infty}} \Delta \sum_{|k| > \frac{A}{\Delta}} |f(k\Delta)| = 0.$$

(3) For absolutely Riemann-integrable functions the last equality is equivalent to the limit relation

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{h \rightarrow 0} h \sum_{k=-\infty}^{\infty} f(kh).$$

It is known that the case of these functions is rather wide.

6.4.12. COMMENTS.

(1) The main results of this section, including the concepts of a σ -finite subspace of a Loeb measure space and \mathcal{S}_M -integrability, are due to Gordon [145, 146]. Many substantial applications of Loeb measure belong to probability, and in this connection most attention has been focused on studying finite Loeb measures.

Finite Radon measures that are induced by Loeb measures as well as the mapping st were studied in [12] and other publications (see the survey [69] and the book [3]). However, the question of conditions for $j \circ {}^*f|_X$ to be a lifting of f was not addressed.

Another construction of a lifting of the interval $[0, 1]$ may be found in [69]. The treatment of σ -finite Loeb measures is essential for our further aims because they are applied in the next chapter to the study of Haar measures on locally compact abelian groups, and the latter are mostly infinite.

(2) Condition 6.4.9 (2) holds automatically for compactly-supported functions and is superfluous for finite measure spaces. The presence of a hyperapproximant in 6.4.9 (2) is a shortcoming. However, 6.4.9 (2) can be reformulated sometimes (see 6.4.11) in standard terms. Moreover, it often happens that a hyperapproximant can be chosen so that 6.4.9 (2) becomes superfluous.

(3) Returning again to the example 6.4.11, choose L and Δ as the follows: Distinguishing an unlimited hyperreal $\tau \in {}^*\mathbb{R}$, take $\Delta \approx 0$ and $L = \lceil \frac{\tau}{\Delta} \rceil$. As above, the triple (X, j, Δ) is a hyperapproximant to the measure space $(\mathfrak{X}, \Omega, \mu)$, where $\mathfrak{X} := \mathbb{R}$, Ω is the σ -algebra of Lebesgue-measurable sets, and μ is Lebesgue measure. From 4.6.13 it follows that 6.4.10 holds for each absolutely Riemann-integrable almost everywhere continuous function on \mathbb{R} . If f is such a function then it is easy to see that ${}^*f \circ j$ is \mathcal{S}_M -integrable on using the \mathcal{S} -integrability of ${}^*f|_{*[-n,n]} \circ j$ for all $n \in \mathbb{N}$, the equality of 6.4.10, and the closure of $\mathcal{S}_M^\#$ in $\mathcal{L}^\#$ (see the proof of Theorem 6.4.7).

(4) Solovay [452] (also see [195]) introduced the concept of a random real as a real belonging to no negligible Gödel-constructible set in the Gödel sense. He successfully used this concept for proving that various propositions in measure theory are independent of the axioms of ZFC.

Kolmogorov's complexity theory has the analogous concept (due to Martin-Löf [354]) of a random 0–1 sequence as a sequence lying in no negligible Markov-constructible set. Similar concepts of a random element for $[0, 1]$ with Lebesgue measure and an independent sequence of random elements in this case are introduced in [520] where Theorem 6.4.4 was established in this environment. The proof rested on the law of large numbers. For a standard finite measure space Theorem 6.4.4 was proved in [163] by rather different arguments.

6.5. Hyperapproximation of Integral Operators

In this section we consider the possibility of approximating an integral operator by a hyperfinite-rank operator.

6.5.1. Recall that, given a hyperfinite set X , a standard real $p \in [1, \infty]$, and a positive hyperreal $\Delta \in {}^*\mathbb{R}_+$, we let $\mathcal{L}_{p,\Delta}^X$ stand for the internal space of functions $F : X \rightarrow {}^*\mathbb{F}$ (with \mathbb{F} either of the fields \mathbb{R} and \mathbb{C}) under the norm

$$\|F\|_{p,\Delta} := \left(\Delta \sum_{\xi \in X} |F(\xi)|^p \right)^{1/p} \quad (F \in \mathcal{L}_{p,\Delta}^X).$$

This space is hyperfinite-dimensional since $\dim(\mathcal{L}_{p,\Delta}^X) = |X|$. The nonstandard hull $(\mathcal{L}_{p,\Delta}^X)^\#$ of $\mathcal{L}_{p,\Delta}^X$ is an external Banach space which is nonseparable if $|X|$ is an unlimited hypernatural. If $p = 2$ then the norm $\|F\|_{2,\Delta}$ is generated by the inner product $(F, G) := \Delta \sum_{\xi \in X} F(\xi)\overline{G(\xi)}$, and so $(\mathcal{L}_{2,\Delta}^X)^\#$ is a nonseparable Hilbert space. As above, we abbreviate $\mathcal{L}_{p,\Delta}^X$ to \mathcal{L}_p , since this leads to no confusion.

If $(M, S_\Delta^M, \nu_\Delta^M)$ is a σ -finite subspace of the Loeb measure space $(X, S_\Delta, \nu_\Delta)$ then $\mathcal{S}_p(M)^\#$ is a closed subspace of $\mathcal{L}_p^\#$ (see 6.3.12 (3)). By 6.3.14 this subspace is isomorphic to $L_p(\nu_\Delta^M)$. Such an isomorphism is established by sending each function $f \in L_p(\nu_\Delta^M)$ to the coset $F^\#$ of a lifting $F \in \mathcal{S}_p(M)$ of f . Bearing this in mind, we will assume below that $L_p(\nu_\Delta^M) \subset \mathcal{L}_p^\#$.

Suppose now that $(\mathfrak{X}_k, \Omega_k, \mu_k)$ for $k := 1, 2$ are standard σ -finite measure spaces. Furthermore, assume given two Loeb measure spaces $(X_k, S_{\Delta_k}, \nu_{\Delta_k})$, σ -finite subspaces $(M_k, S_{\Delta_k}^{M_k}, \nu_{\Delta_k}^{M_k})$ of these spaces, and $J_{p_k}^{M_k} : L_{p_k}(\mu_k) \rightarrow L_{p_k}(\nu_{\Delta_k}^{M_k}) \subset \mathcal{L}_{p_k}^\#$ for some $p_1, p_2 \in [1, +\infty]$.

Let $\mathcal{A} : L_{p_1}(\mu_1) \rightarrow L_{p_2}(\mu_2)$ be a bounded linear operator. A limited internal operator $A : \mathcal{L}_{p_1} \rightarrow \mathcal{L}_{p_2}$ is a *hyperfinite-rank approximant* or, briefly, *hyperapproximant* to \mathcal{A} provided that, given an arbitrary $f \in \mathcal{L}_{p_1}(\mu_1)$, we have

$\|G - A(F)\|_{p_2, \Delta_2} \approx 0$ whenever $F \in \mathcal{L}_{p_1}$ is a lifting of f and $G \in \mathcal{L}_{p_2}$ is a lifting of $\mathcal{A}(f)$. The expressions like “ A hyperapproximates \mathcal{A} ” are also in common parlance.

If, as in 6.4.8, the embedding j_p^M is induced by some internal mapping $j : X \rightarrow {}^*X$, and ${}^*f \circ j$ serves as a lifting of f then we may view ${}^*f \circ j$ as the *table* of f at the knots of the hyperfinite grid $j(X)$ (the equality of 6.4.10 shows that this is a rather natural standpoint). In this event the operator A hyperapproximates \mathcal{A} if A carries the table of f to some vector infinitely close to the table of the image $\mathcal{A}(f)$ of f under \mathcal{A} .

(1) An internal operator $A : \mathcal{L}_{p_1} \rightarrow \mathcal{L}_{p_2}$ is a hyperapproximant to $\mathcal{A} : L_{p_1}(\mu_1) \rightarrow L_{p_2}(\mu_2)$ if and only if the following diagram commutes:

$$\begin{array}{ccc} L_{p_1}(\mu_1) & \xrightarrow{\mathcal{A}} & L_{p_2}(\mu_2) \\ j_{p_1}^{M_1} \downarrow & & \downarrow j_{p_2}^{M_2} \\ \mathcal{L}_{p_1}^\# & \xrightarrow{A^\#} & \mathcal{L}_{p_2}^\#. \end{array}$$

(2) Assume that the linear span of $\mathfrak{M} \subset L_{p_1}(\mu_1)$ is dense in $L_{p_1}(\mu_1)$. If A hyperapproximates \mathcal{A} for all $f \in \mathfrak{M}$ then A is a hyperapproximant to \mathcal{A} .

◁ If $\mathcal{L}(\mathfrak{M})$ stands for the linear span of \mathfrak{M} then by hypothesis the diagram of (1) commutes on substituting $\mathcal{L}(\mathfrak{M})$ for $L_{p_1}(\mu_1)$. By continuity, the original diagram commutes since $\mathcal{L}(\mathfrak{M})$ is dense in $L_{p_1}(\mu_1)$. ▷

6.5.2. Recall that $\mathcal{A}_k : L_{p_1}(\mu_1) \rightarrow L_{p_2}(\mu_2)$ is an *integral operator with kernel* k provided that $k : \mathfrak{X}_1 \times \mathfrak{X}_2 \rightarrow \mathbb{F}$ is a measurable function such that for every $f \in L_{p_1}(\mu_1)$ the value $g := \mathcal{A}_k(f)$ of \mathcal{A}_k at f is determined from the equality

$$g(s) = \int_{\mathfrak{X}_1} k(s, t) f(t) d\mu_1(t).$$

The fact that \mathcal{A}_k is an integral operator with kernel k may be rewritten as

$$\mathcal{A}_k(f)(s) = \int_{\mathfrak{X}_1} k(s, t) f(t) d\mu_1(t) \quad (f \in L_{p_1}(\mu_1)).$$

Let $\mathcal{A}_k : L_{p_1}(\mu_1) \rightarrow L_{p_2}(\mu_2)$ be an integral operator with kernel k . Then the question naturally arises of whether we may construct a hyperapproximant to \mathcal{A} by using a lifting of the kernel k , if such a lifting exists. We should keep in mind here that, generally speaking,

$$\left(M_1 \times M_2, S_{\Delta_1}^{M_1} \otimes S_{\Delta_2}^{M_2}, \nu_{\Delta_1}^{M_1} \otimes \nu_{\Delta_2}^{M_2} \right) \neq \left(M_1 \times M_2, S_{\Delta_1 \Delta_2}^{M_1 \times M_2}, \nu_{\Delta_1 \Delta_2}^{M_1 \times M_2} \right).$$

(The right side of the above is a σ -finite subspace of the Loeb measure space $(X_1 \times X_2, S_{\Delta_1 \Delta_2}^{X_1 \times X_2}, \nu_{\Delta_1 \Delta_2}^{X_1 \times X_2})$.)

However (see [12]), if M_1 and M_2 are finite Loeb measure spaces (in which case M_1 and M_2 may be assumed internal) then $S_{\Delta_1}^{M_1} \otimes S_{\Delta_2}^{M_2} \subset S_{\Delta_1 \Delta_2}^{M_1 \times M_2}$ and the identity embedding is measure-preserving.

It is easy to see that this result remains true also in the case of σ -finite subspaces of Loeb measure spaces, because each σ -finite subspace is the disjoint union of a countable family of finite Loeb measure spaces.

Thus $L_r(\nu_{\Delta_1}^{M_1} \otimes \nu_{\Delta_2}^{M_2})$ is isometrically embedded in $L_r(\nu_{\Delta_1 \Delta_2}^{M_1 \times M_2})$ for all $r \in [1, \infty)$. Since the embedding $j_r^{M_i} : L_r(\mu) \rightarrow L_r(\nu_{\Delta_i}^{M_i})$, $i := 1, 2$, induces the embedding $j_r^{M_1 \times M_2} : L_r(\mu_1 \otimes \mu_2) \rightarrow L_r(\nu_{\Delta_1}^{M_1} \otimes \nu_{\Delta_2}^{M_2}) \subset L_r(\nu_{\Delta_1 \Delta_2}^{M_1 \times M_2})$, we may conclude that every function $k \in L_r(\mu_1 \otimes \mu_2)$ has a lifting $K \in \mathcal{S}_r(M_1 \times M_2)$.

It stands to reason to find the conditions for a matrix K to determine a hyperapproximant A to \mathcal{A} . We will address this problem for Hilbert–Schmidt operators.

Up to the end of this section we agree that $(\mathfrak{X}, \Omega, \mu)$ is a σ -finite measure space, $(M, S_{\Delta}^M, \nu_{\Delta}^M)$ is a σ -finite subspace of the Loeb measure space $(X, S_{\Delta}, \nu_{\Delta})$. Moreover, we specify the embedding $j_2 : L_r(\mu) \rightarrow L_r(\nu_{\Delta}^M)$.

An integral operator $\mathcal{A}_k : L_2(\mu) \rightarrow L_2(\mu)$ with kernel k is a *Hilbert–Schmidt operator* provided that $k \in L_2(\mu \otimes \mu)$.

The following estimate is well known for every Hilbert–Schmidt operator:

$$\|\mathcal{A}_k\| \leq \left(\int_{\mathfrak{X} \times \mathfrak{X}} |k|^2 d\mu \otimes d\mu \right)^{1/2}.$$

Given an internal function $K \in \mathcal{S}_2(M \times M)$, define the internal operator $A_K : \mathcal{L}_2 \rightarrow \mathcal{L}_2$ by the formula

$$A_K(F)(\xi) = \Delta \sum_{\eta \in X} K(\xi, \eta) F(\eta) \quad (F \in \mathcal{L}_2, \xi \in X).$$

Clearly, the norm of A_K satisfies the following inequality:

$$\|A_K\| \leq \left(\Delta^2 \sum_{\xi, \eta \in X} |K(\xi, \eta)|^2 \right)^{1/2}.$$

6.5.3. Theorem. *If $\mathcal{A}_k : L_2(\mu) \rightarrow L_2(\mu)$ is the Hilbert–Schmidt operator with kernel $k \in L_2(\mu \otimes \mu)$, and $K \in \mathcal{S}_2(M \times M)$ is a lifting of k (that is, K is a lifting of $j_2^{M \times M}(k) \in L_2(\nu_{\Delta_2}^{M \times M})$), then $A_K : \mathcal{L}_2 \rightarrow \mathcal{L}_2$ is a hyperapproximant to \mathcal{A}_k .*

◁ Start with showing the closure in $L_2(\mu \otimes \mu)$ of the set of k 's meeting the claim of the theorem.

To this end, suppose that the theorem holds with $k_n \in L_2(\mu \otimes \mu)$ for all $n \in \mathbb{N}$. Suppose also that $\|k - k_n\|_{L_2(\mu \otimes \mu)} \rightarrow 0$ as $n \rightarrow \infty$. Since $\mathcal{A}_k - \mathcal{A}_{k_n} = \mathcal{A}_{k-k_n}$ by the definition of \mathcal{A}_k , it follows from the estimate for the norm of \mathcal{A}_k in 6.5.2 that $\|\mathcal{A}_k - \mathcal{A}_{k_n}\| \rightarrow 0$ as $n \rightarrow \infty$.

Let $K_n \in \mathcal{S}_2(M \times M)$ be a lifting of k_n and let $K \in \mathcal{S}_2(M \times M)$ be a lifting of k . Then

$$\begin{aligned} \|A_K^\# - A_{K_n}^\#\| &= {}^\circ\|A_K - A_{K_n}\| = {}^\circ\|A_{K-K_n}\| \\ &\leq \left(\Delta^2 \sum_{\xi, \eta \in X} |K(\xi, \eta) - K_n(\xi, \eta)|^2 \right)^{1/2} = \|k - k_n\|_{L_2(\mu \otimes \mu)} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

We now use the fact that the diagram of 6.5.1 (2) commutes for \mathcal{A}_{k_n} and A_{K_n} . Take $f \in L_2(\mu)$. Then

$$j_2(\mathcal{A}_k(f)) = \lim_{n \rightarrow \infty} j_2(\mathcal{A}_{k_n}(f)) = \lim_{n \rightarrow \infty} A_{K_n}^\#(j_2(f)) = A_K^\#(j_2(f)).$$

To complete the proof of the theorem we have to show that the claim holds for the functions k of the shape $\varphi \otimes \psi$, i.e., $\varphi \otimes \psi(s, t) := \varphi(s)\psi(t)$. Since linear combinations of these functions are dense in $L_2(\mu \otimes \mu)$, this will complete the proof in view of 6.5.1 (1).

To this end, put $k := \varphi \otimes \psi$, where $\varphi, \psi \in L_2(\mu)$. Denote some liftings of φ and ψ by $\Phi, \Psi \in \mathcal{S}_2(M)$, respectively. Assume also that $F \in \varphi_2(M)$ is a lifting of a function $f \in L_2(\mu)$ and $G \in \mathcal{S}_2(M)$, a lifting of $\mathcal{A}_k(f)$. Obviously, $\eta \mapsto \Psi(\eta)F(\eta)$ is a lifting of $t \mapsto \psi(t)f(t)$.

By using the Cauchy–Bunyakovskii–Schwarz inequality and the membership $\Psi, F \in \mathcal{S}_2(M)$ it is easy to show that $\Psi F \in \mathcal{S}(M)$, yielding

$$\alpha := \int_{\mathfrak{X}} \psi(t)f(t) d\mu(t) \approx \sum_{\eta \in X} \Psi(\eta)F(\eta) =: \beta.$$

Since $\mathcal{A}_k(f) = \alpha\varphi$, the lifting G of $\mathcal{A}_k(f)$ is equal to $\alpha\Phi$. Also, it follows from the definition of A_K that $A_K(F) = \beta\Phi$. So, $\|G - A_K(F)\|_2 = |\alpha - \beta|\|\Phi\|_2 \approx 0$, which completes the proof. ▷

6.5.4. From 6.5.3 and 6.4.7 we infer the following corollary:

(1) *There is a hyperapproximant to each Hilbert–Schmidt operator in the $L_2(\mu)$ space with μ a σ -finite measure.*

Suppose now that \mathfrak{X} is a separable locally compact space, μ is a Borel measure on \mathfrak{X} , and Ω is the completion of the Borel σ -algebra with respect to μ .

Suppose that (X, j, Δ) is a hyperapproximant of the measure space $(\mathfrak{X}, \Omega, \mu)$ (see 6.4.9 (1)). We consider the space $\mathfrak{X} \times \mathfrak{X}$ with the product topology. Obviously, $\text{nst}(*\mathfrak{X} \times *\mathfrak{X}) = \text{nst}(*\mathfrak{X}) \times \text{nst}(*\mathfrak{X})$ and $\text{st}((\xi, \eta)) = (\text{st}(\xi), \text{st}(\eta))$ ($\xi, \eta \in \text{nst}(*\mathfrak{X})$). This implies that $(X \times X, j \otimes j, \Delta^2)$ is a hyperapproximant to the measure space $(\mathfrak{X} \times \mathfrak{X}, \Omega \otimes \Omega, \mu \otimes \mu)$.

It is an easy matter to check that

(2) If $f : \mathfrak{X} \times \mathfrak{X} \rightarrow \mathbb{R}$ is a bounded $\mu \otimes \mu$ -almost everywhere continuous function then the condition 6.4.9 (2) that f has a rather rapid decay at infinity amounts to the following

$$\begin{aligned} & (\forall B \in *\mathcal{P}(X))(B \subset X - M) \\ & \rightarrow \Delta^2 \left(\sum_{x \in B} \sum_{y \in X} |*f(j(x), j(y))| + \sum_{x \in X} \sum_{y \in B} |*f(j(x), j(y))| \right) \approx 0. \end{aligned}$$

(3) Assume that \mathfrak{X} is a separable locally compact topological space with a Borel measure μ and (X, j, Δ) is a hyperapproximant to the measure space $(\mathfrak{X}, \Omega, \mu)$. If k is each bounded $\mu \otimes \mu$ -almost everywhere continuous function such that $k \in L_2(\mu \otimes \mu)$ and $|k|^2$ satisfies (2), then the integral operator A_K with kernel $K := *k|_{j(X) \times j(X)}$ is a hyperapproximant to \mathcal{A}_k .

6.5.5. We now give a simple sufficient condition for $f : \mathfrak{X} \times \mathfrak{X} \rightarrow \mathbb{R}$ to satisfy 6.5.4 (2).

(1) Assume that $f : \mathfrak{X}^2 \rightarrow \mathbb{R}$ satisfies the inequality

$$|f(x, y)| \leq \varphi_1(x)\varphi_2(y) \quad (x, y \in \mathfrak{X}),$$

with φ_1 and φ_2 some bounded integrable μ -almost everywhere continuous functions of a rather rapid decay at infinity (cf. 6.4.9 (2)). Then f satisfies 6.5.4 (2).

◁ By hypothesis, φ_1 and φ_2 satisfy 6.4.9 (2) and 6.4.10. Thus, if $B \subset X - M$ then

$$\begin{aligned} & \Delta^2 \sum_{x \in B} \sum_{y \in X} |*f(j(x), j(y))| \\ & \leq \Delta \sum_{x \in B} *\varphi_1(j(x)) \Delta \sum_{y \in X} *\varphi_2(j(y)) \approx \int_{\mathfrak{X}} \varphi_2 d\mu \Delta \sum_{x \in B} *\varphi_1(j(x)) \approx 0, \end{aligned}$$

which completes the proof. ▷

This proposition amounts to the following:

(2) Assume that $k \in L_2(\mu \otimes \mu)$ is a bounded almost everywhere continuous function such that $|k|^2$ satisfies 6.5.4 (2) (or the inequality of (1)). If $f \in L_2(\mu)$ is a bounded almost everywhere continuous function of a rather rapid decay at infinity (cf. 6.4.9 (2)) then

$$\Delta \sum_{x \in X} \left| \int_{*\mathfrak{X}} *k(j(x), \eta) *f(\eta) d*\mu(\eta) - \Delta \sum_{y \in X} *k(j(x), j(y)) *f(j(y)) \right|^2 \approx 0.$$

6.5.6. In the particular case of $\mathfrak{X} := \mathbb{R}$ (see 6.4.11), we come to the following corollary.

(1) Let $k \in L_2(\mathbb{R}^2)$ be a bounded almost everywhere continuous function such that $|k|^2$ satisfies the inequality of 6.5.5 (1) for some absolutely integrable functions φ_1 and φ_2 on \mathbb{R} satisfying 6.4.11 (3) (which rephrases a rather rapid decay). Assume that $\Delta \approx 0$ and $L \in *\mathbb{N}$ are such that $L\Delta \approx +\infty$. Then

$$\Delta \sum_{\alpha=-M}^M \left| \int_{-\infty}^{\infty} *k(\alpha\Delta, y) *f(y) dy - \Delta \sum_{\beta=-L}^L *k(\alpha\Delta, \beta\Delta) *f(\beta\Delta) \right|^2 \approx 0$$

for every bounded almost everywhere continuous function $f \in L_2(\mathbb{R})$ satisfying 6.4.11 (3).

We proceed by putting $X := \{k \in *\mathbb{N} : |k| \leq L\}$, define $j : X \rightarrow *\mathbb{R}$ by $j(k) := k\Delta$, and consider $K : X^2 \rightarrow *\mathbb{R}$ such that $K := *k|_{j(X) \times j(X)}$. From (1) it follows now that A_K is a hyperapproximant to \mathcal{A}_k (see 6.5.5 (2)).

(2) If $\tau \approx +\infty$, $\Delta \overset{\tau}{\approx} 0$, and $L := \lceil \frac{\tau}{\Delta} \rceil$, then the relation of (1) holds for all bounded almost everywhere continuous functions $k \in L_2(\mathbb{R}^2)$ and $f \in L_2(\mathbb{R})$.

◁ Arguing by analogy with 4.6.13, infer that

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |k(x, y)|^2 dx dy = \overset{\circ}{\left(\Delta^2 \sum_{\alpha, \beta=-L}^L |*k(\alpha\Delta, \beta\Delta)|^2 \right)}.$$

It is easy to show along the lines of (1) that K belongs to $\mathcal{S}_2(M \times M)$ and is a lifting of k , even if $|k|^2$ does not satisfy 6.5.4 (2). ▷

It is worth noting that (1) and (2) remain true on abstracting \mathbb{R} to \mathbb{R}^n with an arbitrary $n \geq 1$ (in this case $k \in L_2(\mathbb{R}^{2n})$).

6.5.7. We now pass to the standard versions of some of the above results. Clearly, the Nelson algorithm is a natural tool for this matter.

By way of example we translate 6.5.4(1). By Theorem 6.4.7, it is possible to assume that $X \subset {}^* \mathfrak{X}$, and ${}^* f|_X$ in $\mathcal{S}_2(M)$ is a lifting of $f \in L_2(\mu)$. We remark that no similar equality holds in general for lifting $K : X^2 \rightarrow {}^* \mathbb{R}$ in $\mathcal{S}_2(M)$ for a function $k \in L_2(\mu \otimes \mu)$, since (X^2, Δ) can fail to meet the claim of Theorem 6.4.5 in regard to $(\mathfrak{X}^2, \Omega \otimes \Omega, \mu \otimes \mu)$. Nevertheless, such a lifting K exists in view of the remarks on 6.5.2.

Let $(\mathfrak{X}, \Omega, \mu)$ be a σ -finite measure space, and let $k \in L_2(\mu \otimes \mu)$. Then, to each finite set $\Phi \subset L_2(\mu)$ and each $\varepsilon > 0$, there are $X \subset \mathfrak{X}$, $\Delta \in \mathbb{R}$, and $K : X^2 \rightarrow \mathbb{R}$ such that

$$\left| \iint_{\mathfrak{X}^2} |k|^2 d\mu - \Delta^2 \sum_{x,y} |K(x,y)|^2 \right| < \varepsilon,$$

and

$$\Delta \sum_{x \in X} \left| \int_{\mathfrak{X}} k(x, \eta) f(\eta) d\mu(\eta) - \Delta \sum_{y \in X} K(x, y) f(y) \right|^2 < \varepsilon$$

for all $f \in \Phi$.

◁ Note that 6.5.4(1) may be rewritten as follows:

Let $(\mathfrak{X}, \Omega, \mu)$ be a standard σ -finite measure space, and let $k \in L_2(\mu \otimes \mu)$ be standard. Then

$$\begin{aligned} & (\exists^{\text{fin}} X \subset \mathfrak{X})(\exists \Delta \in \mathbb{R}_+)(\exists K : X^2 \rightarrow \mathbb{R})(\forall^{\text{st}} f \in L_2(\mu))(\forall^{\text{st}} 0 < \varepsilon \in \mathbb{R}) \\ & \left(\left| \iint_{\mathfrak{X}^2} |k|^2 d\mu \otimes \mu - \Delta^2 \sum_{x,y \in X} |K(x,y)|^2 \right| < \varepsilon \right. \\ & \left. \wedge \Delta \sum_{x \in X} \left| \int_{\mathfrak{X}} k(x, \eta) f(\eta) d\mu(\eta) - \Delta \sum_{y \in X} K(x, y) f(y) \right|^2 < \varepsilon \right). \end{aligned}$$

Denote the matrix of this formula by \mathscr{W} and proceed by idealization to obtain the equivalent proposition

$$\begin{aligned} & (\forall^{\text{st}} \Phi \in \mathcal{P}_{\text{fin}}(L_2(\mu)))(\forall^{\text{st}} \Xi \in \mathcal{P}_{\text{fin}}(\mathbb{R}_+))(\exists^{\text{fin}} X \subset \mathfrak{X})(\exists \Delta \in \mathbb{R}_+) \\ & (\exists K : X^2 \rightarrow \mathbb{R})(\forall f \in \Phi)(\forall \varepsilon \in \Xi) \mathscr{W}. \end{aligned}$$

Since $(\mathfrak{X}, \Omega, \mu)$ and k are standard, we may remove the superscript st from the first two quantifiers by transfer. Passing from Ξ to $\varepsilon := \min(\Xi)$ enables us to replace the second quantifier by $(\forall \varepsilon \in \mathbb{R}_+)$. Thus, we may eliminate the last quantifier. ▷

A more careful inspection of the above translation shows that while the non-standard entities X , Δ , and K in 6.5.4(1) depend only on the elements of $L_2(\mu)$ and $L_2(\mu \otimes \mu)$, which are cosets, rather than on their constituents, and while the $k \in L_2(\mu \otimes \mu)$ is regarded as a coset (i.e., X , Δ , and K do not depend on the choice of a constituent in the coset k). Observe also that Φ in this formula is a finite set of square-integrable functions.

In fact, Proposition 6.5.7 is not an exact analog of 6.5.4(1), since the first inequality of this proposition is only one of the consequences of the fact that K in $\mathcal{S}(M)$ is a lifting of k . Since the phrase “ K in $\mathcal{S}(M)$ is a lifting of k ” manipulates with strictly external entities, it cannot be translated into the language of IST, but it might give rise to various internal consequences strengthening 6.5.7.

6.5.8. We now consider a standard version of 6.5.6(2). Denote by \mathcal{H}_m the space of almost everywhere continuous bounded functions in $L_2(\mathbb{R}^m)$ (we imply Lebesgue measure on \mathbb{R}^m).

If $\Phi \subset \mathcal{H}_1$ and $\mathcal{K} \subset \mathcal{H}_2$ are finite sets of functions then to each $n \in \mathbb{N}$ there are $L \in \mathbb{N}$ and $\Delta \in \mathbb{R}_+$ satisfying $L\Delta > n$, $\Delta < n^{-1}$, and

$$\Delta \sum_{\alpha=-L}^L \left| \int_{-\infty}^{\infty} k(\alpha\Delta, \eta) f(\eta) d\eta - \Delta \sum_{\beta=-L}^L k(\alpha\Delta, \beta\Delta) f(\beta\Delta) \right|^2 < \frac{1}{n}$$

for all $f \in \Phi$ and $k \in \mathcal{K}$.

◁ Rewrite 6.5.6(2) as follows:

$$\begin{aligned} & (\exists L \in \mathbb{N})(\exists \Delta \in \mathbb{R}_+) \left(L\Delta \approx +\infty \wedge \Delta \approx 0 \wedge (\forall^{\text{st}} f \in \mathcal{H}_1)(\forall^{\text{st}} k \in \mathcal{H}_2) \right. \\ & \left. \left(\Delta \sum_{\alpha=-M}^M \left| \int_{-\infty}^{\infty} *k(\alpha\Delta, y) *f(y) dy - \Delta \sum_{\beta=-L}^L *k(\alpha\Delta, \beta\Delta) *f(\beta\Delta) \right|^2 \approx 0 \right) \right). \end{aligned}$$

Deciphering the symbol \approx , obtain

$$\begin{aligned} & (\exists L \in \mathbb{N})(\exists \Delta \in \mathbb{R}_+)(\forall^{\text{st}} n \in \mathbb{N})(\forall^{\text{st}} f \in \mathcal{H}_1)(\forall^{\text{st}} k \in \mathcal{H}_2) \\ & \left(L\Delta > n \wedge \Delta < n^{-1} \right. \\ & \left. \wedge \left(\Delta \sum_{\alpha=-L}^L \left| \int_{-\infty}^{\infty} k(\alpha\Delta, \eta) f(\eta) d\eta - \Delta \sum_{\beta=-L}^L k(\alpha\Delta, \beta\Delta) f(\beta\Delta) \right|^2 < \frac{1}{n} \right) \right). \end{aligned}$$

In much the same way as before, it is easy to eliminate the predicate st. ▷

6.5.9. To formulate a standard version of 6.5.6 (1) is even simpler. To this end, denote by $\widetilde{\mathcal{H}}_1$ the set of functions $f \in \mathcal{H}_1$ with $|f|^2$ satisfying 6.4.11 (2), and let $\widetilde{\mathcal{H}}_2$ stand for the set comprising $k \in \mathcal{H}_2$ such that $|k|^2$ enjoys the inequality of 6.5.5 (1) for some φ_1 and φ_2 satisfying 6.4.11 (2). (The set $\widetilde{\mathcal{H}}_m$ is defined similarly.)

If $f \in \widetilde{\mathcal{H}}_1$ and $k \in \widetilde{\mathcal{H}}_2$ then

$$\lim_{\substack{\Delta \rightarrow 0 \\ L\Delta \rightarrow \infty}} \Delta \sum_{\alpha=-L}^L \left| \int_{-\infty}^{\infty} k(\alpha\Delta, \eta) f(\eta) d\eta - \Delta \sum_{\beta=-L}^L k(\alpha\Delta, \beta\Delta) f(\beta\Delta) \right|^2 = 0.$$

◁ The relevant formula of 6.5.6 (1) may be rewritten in IST as

$$\begin{aligned} & (\forall L \in \mathbb{N})(\forall \Delta \in \mathbb{R}_+) \left(L\Delta \approx +\infty \wedge \Delta \approx 0 \rightarrow (\forall^{\text{st}} f \in \widetilde{\mathcal{H}}_1)(\forall^{\text{st}} k \in \widetilde{\mathcal{H}}_2) \right. \\ & \left. \left(\Delta \sum_{\alpha=-L}^L \left| \int_{-\infty}^{\infty} k(\alpha\Delta, \eta) f(\eta) d\eta - \Delta \sum_{\beta=-L}^L k(\alpha\Delta, \beta\Delta) f(\beta\Delta) \right|^2 \approx 0 \right) \right). \end{aligned}$$

Since $a \approx 0 \leftrightarrow (\forall^{\text{st}} n \in \mathbb{N})(|a| < 1/n)$, by predicate logic arrive at the following equivalent formula

$$\begin{aligned} & (\forall^{\text{st}} f \in \widetilde{\mathcal{H}}_1)(\forall^{\text{st}} k \in \widetilde{\mathcal{H}}_2)(\forall^{\text{st}} n \in \mathbb{N})(\forall L \in \mathbb{N})(\forall \Delta \in \mathbb{R}_+) \\ & (L\Delta \approx +\infty \wedge \Delta \approx 0 \rightarrow \mathcal{W}), \end{aligned}$$

with \mathcal{W} denoting the inequality of 6.5.5 (1).

Writing the predicates $L\Delta \approx +\infty$ and $\Delta \approx 0$ in IST, by idealization and transfer arrive at the proposition

$$\begin{aligned} & (\forall f \in \widetilde{\mathcal{H}}_1)(\forall k \in \widetilde{\mathcal{H}}_2)(\forall n \in \mathbb{N})(\exists^{\text{fin}} \mathcal{N} \subset \mathbb{N})(\forall L \in \mathbb{N})(\forall \Delta \in \mathbb{R}_+) \\ & (\forall m \in \mathcal{N})((\Delta < 1/m \wedge L\Delta > m) \rightarrow \mathcal{W}), \end{aligned}$$

which completes the proof. ▷

6.5.10. We consider a standard version of the definition of a hyperapproximant to an arbitrary bounded linear operator $\mathcal{A} : L_p(\mathbb{R}) \rightarrow L_q(\mathbb{R})$, where $p, q \geq 1$.

Denote by $\widetilde{\mathcal{H}}^{(r)}$ the space of bounded almost everywhere continuous functions $f \in L_r(\mathbb{R})$ such that $|f|^2$ satisfies 6.4.11 (2). Note that $\widetilde{\mathcal{H}}^{(r)}$ is dense in $L_r(\mathbb{R})$ ($r \geq 1$). We also assume that \mathcal{A} satisfies the condition:

(1) $\{f \in \widetilde{\mathcal{H}}^{(p)} : \mathcal{A}(f) \in \widetilde{\mathcal{H}}^{(q)}\}$ is dense in $L_p(\mathbb{R})$.

Let $\mathbf{T} := \{T_{L,\Delta} : L \in \mathbb{N}, \Delta \in \mathbb{R}_+\}$, with $T_{L,\Delta} := (t_{\alpha,\beta})_{\alpha,\beta=-L}^L$, be a matrix pencil depending on two parameters and uniformly bounded in norm. Each of the matrices $T_{L,\Delta}$ has size $N \times N$, with $N := 2L + 1$.

(2) A matrix pencil \mathbf{T} approximates a bounded linear operator $\mathcal{A} : L_p(\mathbb{R}) \rightarrow L_q(\mathbb{R})$ satisfying (1) provided that

$$\lim_{\substack{\Delta \rightarrow 0 \\ L\Delta \rightarrow \infty}} \Delta \sum_{\alpha=-L}^L \left| \mathcal{A}(f)(\alpha\Delta) - \sum_{\beta=-L}^L t_{\alpha\beta} f(\beta\Delta) \right|^q = 0$$

for all $f \in \widetilde{\mathcal{H}}^{(p)} \cap \mathcal{A}^{-1}(\widetilde{\mathcal{H}}^{(q)})$.

(3) Assume that \mathbf{T} is a matrix pencil satisfying the above-stated conditions. Then \mathbf{T} approximates a bounded operator $\mathcal{A} : L_p(\mathbb{R}) \rightarrow L_q(\mathbb{R})$ if and only if for all $L \in {}^*\mathbb{N}$ and $\Delta \in {}^*\mathbb{R}_+$ with $\Delta \approx 0$ and $L\Delta \approx +\infty$ the operator ${}^*T_{L,\Delta} : \mathcal{L}_{\Delta,p}^X \rightarrow \mathcal{L}_{\Delta,q}^X$, with $X := \{-L, \dots, L\}$, is a hyperapproximant to \mathcal{A} .

◁ This is immediate from 6.5.10 (1) and the infinitesimal limit test (cf. 2.3.1). ▷

(4) Assume that the equality of (2) holds for all functions f belonging to some set $\mathfrak{M} \subset \widetilde{\mathcal{H}}^{(p)} \cap \mathcal{A}^{-1}(\widetilde{\mathcal{H}}^{(q)})$, whose linear span is dense in $L_p(\mathbb{R})$. Then this equality holds for all $f \in \widetilde{\mathcal{H}}^{(p)} \cap \mathcal{A}^{-1}(\widetilde{\mathcal{H}}^{(q)})$.

◁ This follows from 6.5.1 (1). ▷

6.5.11. COMMENTS.

(1) The results of this section are due to Gordon and published in [140, 146]. Our presentation follows [146].

(2) The proposition “ \mathcal{A} has a hyperapproximant” is expressible in the language of IST. Therefore, we may rephrase this proposition in the standard mathematical terms by the Nelson algorithm. In full generality, such a reformulation is rather bulky. However, it means essentially that there are a sequence of finite-dimensional spaces and a sequence of finite-rank operators such that the relevant sequences of finite grids (comprising the knots of tables of the consecutive spaces), the sequence of meshsizes Δ such that the table of a function at the knots of each grid is a vector in the respective space, the integral of a function f is approximated with the sum of the values of f at the knots with meshsize Δ , and the values of the finite-rank operator at the table of f converge to the table of $\mathcal{A}(f)$; cf. 6.5.10 (2).

(3) The results, analogous to Theorem 6.5.3 and its corollaries, can be abstracted to some other classes of integral operators on imposing the conditions on k under which the integral operator with kernel k is bounded from L_p to L_q (for example, see [227]).

(4) The infinitesimal definition of hyperapproximation in 6.5.1 is essentially more general than 6.5.10 (2) even in the case of the $L_p(\mathbb{R})$ spaces, since the former abstains in general from presuming the existence of a standard matrix pencil \mathbf{T} satisfying the conditions prior to 6.5.10 (2). The distinction between these definitions transpires on comparison between 6.5.8 and 6.5.9. The latter may be refined by a complete translation of 6.5.6 (2) into the standard language on considering that $\Delta \overset{\tau}{\approx} 0$. However, this refinement involves bulky and bizarre formulas.

(5) In the sequel we will need the case in which the table of f in the approximation of \mathcal{A} is taken on a uniform grid with meshsize Δ , whereas the table of $\mathcal{A}(f)$ is taken with meshsize Δ_1 . Of course, $\Delta_1 \rightarrow 0$ and $L\Delta_1 \rightarrow \infty$ (for instance, put $\Delta_1 := ((2L + 1)\Delta)^{-1}$). The general definition of 6.5.10 obviously encompasses this case. Therefore, 6.4.9 (2) and 6.5.10 (3, 4) remain valid after slight changes.

(6) It is possible to abstract 6.4.9 (2) and 6.5.10 (3, 4) to the case in which we hyperapproximate an operator $\mathcal{A} : L_p(\mu) \rightarrow L_q(\mu)$, with μ a σ -finite measure on a separable locally compact topological space \mathfrak{X} , by using hyperapproximation the corresponding measure space on \mathfrak{X} (see 6.5.4 (3)).

Here it is necessary to give a standard version of the definition of hyperapproximant to a measure space as some collection of a family X of finite sets, mappings $j : X \rightarrow \mathfrak{X}$, and reals Δ satisfying appropriate conditions. As was noted, the main difficulties on this way are encountered in translating 6.4.9 (2) and 6.5.4 (2) to the standard language.

We will return to similar questions in the next chapter when addressing hyperapproximation of locally compact abelian groups.

6.6. Pseudointegral Operators and Random Loeb Measures

We need a somewhat stronger version of the theorem of 6.4.6. Namely, it is desirable now that the integral of each integrable function be approximated by the sum over a hyperfinite set to within a given infinitesimal ε .

6.6.1. Theorem. *Let $(\mathcal{X}, \Omega, \mu)$ be a standard σ -finite measure space and let ε be a strictly positive infinitesimal. Then there are an internal hyperfinite set $X \subseteq {}^*\mathcal{X}$ and a hyperreal $\Delta \in {}^*\mathbb{R}$ such that*

$$\left| \int_{\mathcal{X}} f d\mu - \Delta \sum_{\xi \in X} {}^*f(\xi) \right| < \varepsilon$$

for all $f \in L_1(\mu)$.

◁ Let k be some ε -infinite integer. Then $({}^*\mathcal{X}_k, {}^*\Omega_k, {}^*\mu_k)$ satisfies the conditions of Theorem 6.4.4, and so there is an internal hyperfinite set $X \subseteq {}^*\mathcal{X}_k$ such that

$$\int_{{}^*\mathcal{X}_k} h \, d\mu_k - \frac{\mu_k({}^*\mathcal{X}_k)}{|X|} \sum_{\xi \in X} h(\xi)$$

is k -infinite for every k -standard integrable function $h : {}^*\mathcal{X}_k \rightarrow {}^*\mathbb{R}$. In particular, if $f \in L_1(\mu)$ then *f_k is a k -standard integrable function on ${}^*\mathcal{X}_k$; therefore, the hyperreal

$$\int_{{}^*\mathcal{X}_k} {}^*f_k \, d\mu_k - \frac{\mu_k({}^*\mathcal{X}_k)}{|X|} \sum_{\xi \in X} {}^*f_k(\xi)$$

is k -infinitesimal and, consequently, it is ε -infinitesimal.

Furthermore, from the equality

$$\int_{\mathcal{X}} f \, d\mu = \lim_{n \rightarrow \infty} \int_{\mathcal{X}_n} f_n \, d\mu_n$$

it follows that $\int_{\mathcal{X}} f \, d\mu - \int_{{}^*\mathcal{X}_k} {}^*f_k \, d\mu_k$ is ε -infinitesimal. Since ${}^*f_k|_X = {}^*f|_X$; therefore, the difference

$$\int_{\mathcal{X}} f \, d\mu - \frac{\mu_k({}^*\mathcal{X}_k)}{|X|} \sum_{\xi \in X} {}^*f_k(\xi)$$

is ε -infinitesimal too and so its modulus is at most ε . ▷

6.6.2. In the context of Theorem 6.6.1 we say that the couple (X, Δ) *approximates* μ to within ε .

The proof of Theorem 6.6.1 proceeds exclusively by transfer and saturation, we may replace standardness by relative standardness.

More precisely, if τ is a distinguished internal set, $(\mathcal{X}, \Omega, \mu)$ is a τ -standard σ -finite measure space μ , and ε is a positive τ -infinitesimal; then there are an internal hyperfinite set $X \subseteq {}^*\mathcal{X}$ and a hyperreal $\Delta \in {}^*\mathbb{R}$ such that

$$\left| \int_{{}^*\mathcal{X}} F \, d{}^*\mu - \Delta \sum_{\xi \in X} F(\xi) \right| < \varepsilon$$

for every τ -standard integrable function F . In particular, this remains valid for every infinitesimal ε .

6.6.3. We assume that \mathcal{X} is an arbitrary set, \mathcal{A} is some σ -algebra of subsets of \mathcal{X} , and $(\lambda_y)_{y \in \mathcal{Y}}$ is a standard family of σ -finite measures on \mathcal{A} . The family $(\lambda_y)_{y \in \mathcal{Y}}$ may be viewed as the function $\lambda : \mathcal{A} \times \mathcal{Y} \rightarrow \overline{\mathbb{R}}$ satisfying $\lambda_y(A) := \lambda(A, y)$ for all $y \in \mathcal{Y}$ and $A \in \mathcal{A}$. Let \mathcal{L}_1 stand for the set of all measurable functions on \mathcal{X} measurable with respect to each measure of the family $(\lambda_y)_{y \in \mathcal{Y}}$. The family $(\lambda_y)_{y \in \mathcal{Y}}$ generates the pseudointegral operator T on \mathcal{L}_1 by the rule: if $f \in \mathcal{L}_1$ then Tf is the function from \mathcal{Y} to \mathbb{R} , acting as

$$(Tf)(y) = \int_{\mathcal{X}} f d\lambda_y \quad (f \in \mathcal{L}_1).$$

In the next subsection we will show that we may approximate a pseudointegral operator with a matrix to within infinitesimal.

Denote the vector space of all real functions on \mathcal{Y} by $\mathcal{F}(\mathcal{Y})$. Given a hyperfinite collection $X := (x_1, \dots, x_n) \subseteq {}^*\mathcal{X}$, we let the symbol π_X stand for the “projection” from \mathcal{L}_1 to ${}^*\mathbb{R}^n$ which sends each function $f \in \mathcal{L}_1$ to the table of f , i.e., the vector $({}^*f(x_1), \dots, {}^*f(x_n))$.

6.6.4. Theorem. In $\mathcal{Y} \subseteq Y$ there are finite collections of elements $X := (x_1, \dots, x_n) \subseteq {}^*\mathcal{X}$ and $Y := (y_1, \dots, y_m) \subseteq {}^*\mathcal{Y}$ together with a matrix $A = (a_{ij})$ of size $n \times m$ such that $\pi_Y({}^*Tf) \approx A\pi_X({}^*f)$ for every function $f \in \mathcal{L}_1$; i.e.,

$$\int_{{}^*\mathcal{X}} {}^*f d{}^*\lambda_{y_i} \approx \sum_{j=1}^n a_{ij} {}^*f(x_j) \quad (i := 1, \dots, m).$$

In other words, the following diagram commutes to within infinitesimal:

$$\begin{array}{ccc} \mathcal{L}_1(\mathcal{X}) & \xrightarrow{T} & \mathcal{F}(\mathcal{Y}) \\ \pi_X \downarrow & & \pi_Y \downarrow \\ {}^*\mathbb{R}^n & \xrightarrow{A} & {}^*\mathbb{R}^m \end{array}$$

◁ Take a strictly positive infinitesimal ε . For all $\lambda \in \mathcal{Y}$ from Theorem 6.6.1 it follows that there are a hyperfinite sequence $X(y)$ of elements in ${}^*\mathcal{X}$ and a positive hyperreal $\Delta(y)$ such that $|\int_{{}^*\mathcal{X}} f d\lambda_y - \Delta(y) \sum_{X(y)} f| \leq \varepsilon$ for every standard function $f \in \mathcal{L}_1$. This gives rise to the functions

$$X : \mathcal{Y} \rightarrow \bigcup_{n \in {}^*\mathbb{N}} {}^*\mathcal{X}^n \quad \text{and} \quad \Delta : \mathcal{Y} \rightarrow {}^*\mathbb{R}.$$

By extension we may assume that these functions are given on ${}^*\mathcal{Y}$. If $y \in {}^*\mathcal{Y}$ and $F \in {}^*\mathcal{L}_1$ then agree to say that $\Phi(y, F)$ holds whenever

$$\left| \int_{{}^*\mathcal{X}} F d^*\lambda_y - \Delta(y) \sum_{X(y)} F \right| \leq \varepsilon.$$

Given $y \in \mathcal{Y}$, consider the internal set $B_y := \{F \in {}^*\mathcal{L}_1 : \Phi(y, F)\}$. Note that ${}^*f \in B_y$ for all $f \in \mathcal{L}_1$ and $y \in \mathcal{Y}$. Furthermore, to each finite collection $(y_1, f_1), \dots, (y_k, f_k) \in \mathcal{Y} \times \mathcal{L}_1$ there is an internal set ${}^*\mathcal{L}_1$ containing all f_m and lying in each B_{y_m} , $m := 1, \dots, k$. As such a set we may take $B_{y_1} \cap \dots \cap B_{y_k}$. By saturation it follows now that there is an internal set B satisfying $f \in B \subseteq F_y$ for all $(y, f) \in \mathcal{Y} \times \mathcal{L}_1$. In particular, $\Phi(y, F)$ holds for all $y \in \mathcal{Y}$ and $F \in B$.

Let Y_0 stand for the set of $y \in {}^*\mathcal{Y}$ such that $\Phi(y, F)$ holds for all $F \in B$. Clearly, Y_0 is an internal set including \mathcal{Y} . On the other hand, there is a hyperfinite set $Y_1 \subseteq {}^*\mathcal{Y}$ including \mathcal{Y} . Put $Y := Y_0 \cap Y_1$. Then Y is a hyperfinite internal subset of ${}^*\mathcal{Y}$ including \mathcal{Y} . Moreover, $\Phi(y, F)$ holds for all $y \in Y$ and $F \in B$. In particular, $\Phi(y, {}^*f)$ holds for all $y \in Y$ and $f \in \mathcal{L}_1$. Order Y somehow to obtain the tuple (y_1, \dots, y_m) . As X we take the concatenation of the tuples X_{y_1}, \dots, X_{y_m} , i.e., the tuple consisting of the tuples X_{y_1}, \dots, X_{y_m} located consecutively. For definiteness, assume that $X := (x_1, \dots, x_n)$ and the elements of the tuple X_{y_i} take the places from s_{i+1} to s_{i+m} in X . Define the ‘‘step’’ matrix A of size $m \times n$ as follows:

$$a_{ij} = \begin{cases} \Delta_{y_i}, & \text{in case } s_i < j \leq s_{i+1}, \\ 0 & \text{otherwise.} \end{cases}$$

Given $f \in \mathcal{L}_1$, for all $i := 1, \dots, m$ we then see that

$$\left| \int_{{}^*\mathcal{X}} {}^*f d^*\lambda_{y_i} - \sum_{j=1}^n a_{ij} {}^*f(x_j) \right| = \left| \int_{{}^*\mathcal{X}} {}^*f d^*\lambda_{y_i} - \Delta(y_i) \sum_{X(y_i)} {}^*f \right| \leq \varepsilon,$$

i.e., $\int_{{}^*\mathcal{X}} {}^*f d^*\lambda_{y_i} \approx \sum_{j=1}^n a_{ij} {}^*f(x_j)$. The proof is complete. \triangleright

6.6.5. We now consider a particular instance of a pseudointegral operator: an integral operator with kernel $K(x, y)$. Suppose that, in the context of Theorem 6.6.4 we have some σ -finite measure on the σ -algebra \mathcal{A} such that, for all $y \in \mathcal{Y}$, the measure λ_y is absolutely continuous with respect to μ with the density $K_y := K(\cdot, y) \in \mathcal{L}_\infty(\mu)$. Then

$$(Tf)(y) = \int_{\mathcal{X}} f d\lambda_y = \int_{\mathcal{X}} f K_y d\mu.$$

In this event we may slightly strengthen Theorem 6.6.4. Namely, we may apply Theorem 6.6.1 to obtain a hyperfinite set X and a hyperreal Δ so as to approximate μ up to ε . As the matrix we may now take the table of the kernel at the knots of the grid $X \times Y$ with weight Δ .

6.6.6. Theorem. *Suppose that the couple (X, Δ) approximates μ . Then there is a finite tuple $Y = (y_1, \dots, y_m) \subseteq {}^*\mathcal{Y}$ such that $\mathcal{Y} \subseteq Y$ and $\pi_Y({}^*Tf) \approx A\pi_X({}^*f)$, with $a_{ij} := \Delta K(x_i, y_j)$.*

◁ The proof proceeds along the lines of that of Theorem 6.6.4.

Let $\Psi(y, F)$ stand for the internal formula

$$\left| \int F d{}^*\lambda_y - \Delta \sum_X {}^*f {}^*K_y \right| \leq \varepsilon.$$

Given $f \in \mathcal{L}_1$ and $y \in \mathcal{Y}$, note that

$$\left| \int_{\mathcal{X}} f d\lambda_y - \Delta \sum_X {}^*f {}^*K_y \right| = \left| \int_{\mathcal{X}} f K_y d\mu - \Delta \sum_X {}^*f {}^*K_y \right| \leq \varepsilon,$$

i.e., $\Psi(y, {}^*f)$ for all $y \in \mathcal{Y}$ and $f \in \mathcal{L}_1$.

Put $C_y := \{F \in \mathcal{L}_1 : \Psi(y, F)\}$. Then there is an internal set $C \subseteq {}^*\mathcal{L}_1$ such that $\Psi(y, F)$ holds for all $y \in \mathcal{Y}$ and $F \in C$ and, moreover, $\mathcal{L}_1 \subseteq C$. Also, there is a hyperfinite internal tuple $Y := (y_1, \dots, y_m)$ such that $\mathcal{Y} \subseteq Y$ and $\Psi(y, {}^*f)$ holds for all $y \in Y$ and $f \in \mathcal{L}_1$. Then

$$\int_{{}^*\mathcal{X}} {}^*f d\lambda_{y_i} \approx \Delta \sum_X f K_{y_i} = \sum_{j=1}^n a_{ij} f(x_j)$$

for all $i := 1, \dots, m$. ▷

6.6.7. It is worth observing that the proofs of Theorems 6.6.4 and 6.6.6 give a somewhat stronger result: The matrix A of these theorems approximates T to within a given infinitesimal ε . Moreover, since the hyperfinite Y includes \mathcal{Y} ; therefore, the table of a standard function g on \mathcal{Y} is completely determined from the table of *g on Y . This implies in particular that the projection π_Y preserves the supremum of a bounded standard function on \mathcal{Y} , i.e., $\sup_{y \in \mathcal{Y}} g(y) = {}^\circ \max \pi_Y(g)$. On the other hand, the projection π_X in Theorem 6.6.6 preserves the L_1 -norm of f in \mathcal{L}_1 , i.e., $\int_{\mathcal{X}} f d\mu = {}^\circ(\Delta \sum_X {}^*f)$. Consequently, if we furnish ${}^*\mathbb{R}^m$ with the sup-norm then Theorem 6.6.6 implies that to each couple (X, Δ) approximating μ to within ε , there is a hyperfinite Y such that $\mathcal{Y} \subseteq Y \subseteq {}^*\mathcal{Y}$ and $\|\pi_Y({}^*Tf) - A\pi_X({}^*f)\| \leq \varepsilon$ for all $f \in \mathcal{L}_1$.

The next theorem clarifies Theorem 6.6.6. As in the latter, T stands for an integral operator from L_1 to $\mathcal{F}(\mathcal{Y})$ with kernel K , and $\varepsilon > 0$ is a given infinitesimal.

6.6.8. Theorem. *To each hyperfinite $Y \subseteq {}^*\mathcal{Y}$ there is a couple (X, Δ) approximating μ and satisfying $\|\pi_Y({}^*Tf) - A\pi_X({}^*f)\| \leq \varepsilon$ for all $f \in \mathcal{L}_1$ (the matrix A is the same as in Theorem 6.6.4).*

\triangleleft Put $Y := \{y_1, \dots, y_m\}$. If $i := 1, \dots, m$ then K_{y_i} belongs to the set $\{K_{y_1}, \dots, K_{y_m}\}$ and so this set is Y -standard. The space $(\mathcal{X}, \mathcal{A}, \mu)$ is standard and so Y -standard. By 6.6.2 there are a hyperfinite family $X \subseteq {}^*\mathcal{X}$ and a positive hyperreal Δ satisfying

$$\left| \int_{{}^*\mathcal{X}} F d^*\mu - \Delta \sum_X F \right| \leq \varepsilon$$

for every Y -standard integrable function F .

If $f \in \mathcal{L}_1$ then *fK_y is a Y -standard integrable function; consequently,

$$\left| \int_{{}^*\mathcal{X}} {}^*f d^*\lambda_y - \Delta \sum_X fK_y \right| = \left| \int_{{}^*\mathcal{X}} {}^*fK_y d^*\mu - \Delta \sum_X fK_y \right| \leq \varepsilon,$$

which completes the proof. \triangleright

6.6.9. We now abstract the Loeb construction to a random measure which is defined in the following environment: Assume that X is an arbitrary set, \mathcal{A} is an algebra of subsets of X , and (Y, \mathcal{B}, ν) is a measure space with ν finitely additive. A *random measure* (*random finitely additive measure*) is a function $\lambda : \mathcal{A} \times Y \rightarrow \overline{\mathbb{R}}$ satisfying the two conditions:

(1) $\lambda_A := \lambda(A, \cdot) : Y \rightarrow \overline{\mathbb{R}}$ is \mathcal{B} -measurable for all $A \in \mathcal{A}$;

(2) $\lambda_y := \lambda(\cdot, y)$ is a measure (finitely additive measure) on \mathcal{A} for almost all $y \in Y$.

In the sequel, (X, \mathcal{A}) and (Y, \mathcal{B}, ν) are internal sets, with λ an internal finitely additive random measure on $\mathcal{A} \times Y$.

By definition, Y contains some subset Y_0 of full measure such that λ_y is a finitely additive measure for all $y \in Y_0$. Assume further that $(Y, \mathcal{B}_L, \nu_L)$ stands for the Loeb measure space for (Y, \mathcal{B}, ν) . Given λ_y ($y \in Y_0$), consider the corresponding Loeb measure $(\lambda_y)_L$. Observe that the domain of $(\lambda_y)_L$ includes $\sigma(\mathcal{A})$ by the definition of Loeb measure. Define the function $\lambda^L : \sigma(\mathcal{A}) \times Y \rightarrow \overline{\mathbb{R}}$ as follows: Put $\lambda^L(A, y) := (\lambda_y)_L(A)$ for $A \in \sigma(\mathcal{A})$ and $y \in Y_0$, while defining λ^L on $Y - Y_0$ arbitrarily.

6.6.10. Theorem. *The function λ^L constructed above is a random measure with respect to the spaces $(X, \sigma(\mathcal{A}))$ and $(Y, \mathcal{B}_L, \nu_L)$.*

\triangleleft By definition $\lambda_y^L := (\lambda_y)_L$ is a countable additive measure for ν_L -almost all $y \in Y$.

It suffices to show that λ_A^L is \mathcal{B}_L -measurable for all $A \in \sigma(\mathcal{A})$. To this end, denote by M the set of $A \in \sigma(\mathcal{A})$ such that λ_A^L is a \mathcal{B}_L -measurable function.

Clearly, $\mathcal{A} \subseteq M$. Indeed, given $A \in \mathcal{A}$, note that $\lambda_A^L(y) = \lambda_y^L(A) = \circ\lambda_y(A) = \circ\lambda_A(y)$ for all $y \in Y_0$. Thus, λ_A is a lifting of λ_A^L , it follows that the function λ_A^L is \mathcal{B}_L -measurable by Theorem 6.3.6.

We now let $(A_n)_{n \in \mathbb{N}}$ be a monotone sequence of sets in M , and put $A := \lim_{n \rightarrow \infty} A_n$. Then $A \in \sigma(\mathcal{A})$. Since $\lambda_y^L(A) = \lim_{n \rightarrow \infty} \lambda_y^L(A_n)$ for all $y \in Y_0$, the function λ_A^L is \mathcal{B}_L -measurable as the almost everywhere limit of the sequence of \mathcal{B}_L -measurable functions $(\lambda_{A_n}^L)$.

Thus, M is a monotone class. Since every class containing some algebra contains the σ -algebra generated by this algebra; therefore, $M = \sigma(\mathcal{A})$. \triangleright

6.6.11. We have constructed a random Loeb measure on the σ -algebra $\sigma(\mathcal{A})$. Recall that by construction a Loeb measure is defined on the σ -algebra generated by the original algebra after that it is extended to the completion of this σ -algebra. The next example demonstrates that such an extension is impossible in general in the case of a random Loeb measure. Even if the completion of $\sigma(A)$ with respect to each measure λ_y^L is the same for all $y \in Y$, the extension of λ^L on this completion can fail to be a random measure.

Let Y be a hyperfinite set, and let ν be the uniform probability measure on the algebra \mathcal{B} of all internal subsets of Y . It is known in this case that \mathcal{B}_L differs from the powerset $\mathcal{P}(Y)$. Let N be a \mathcal{B}_L -nonmeasurable subset Y . Assign $X := Y$, $\mathcal{A} := \mathcal{B}$, and

$$\lambda(A, y) := \begin{cases} 1, & \text{if } y \in A, \\ 0 & \text{otherwise} \end{cases}$$

for all $y \in Y$.

Obviously, λ is a random measure with respect to (X, \mathcal{A}) and (Y, \mathcal{B}, ν) ; furthermore, for each $y \in Y$ the corresponding Loeb measure $(\lambda_y)_L$ is defined on the whole $\mathcal{P}(X)$. Therefore, we may naturally to extend λ^L onto $\mathcal{P}(X) \times Y$ by putting $\lambda^L(A, y) := (\lambda_y)_L(A)$. However, the function λ^L on $\mathcal{P}(X) \times Y$ is not a random measure since $\lambda_N^L = \chi_N$ is not \mathcal{B}_L -measurable.

6.6.12. We further show that each random measure may be treated as a vector measure and the above constructed Loeb extension of a random measure is in a sense the Loeb extension of a vector measure.

(1) Recall that the nonstandard hull of $V^\#$ of an internal normed vector space V is the quotient space V_1/V_2 where $V_1 := \text{ltd}(V)$ and $V_2 := \mu(V)$; cf. 6.1.1. We will proceed by analogy.

Assume that F is an internal finitely additive V -valued measure on an internal measurable space (X, \mathcal{A}) . Assume also that the image of F lies in some subspace V_1 . Then the function $F^\# : \mathcal{A} \rightarrow V^\#$, acting by the rule $F^\#(A) := F(V)^\#$, is countably

additive, i.e., a measure on \mathcal{A} . It would be natural to call a *vector Loeb measure* an extension of $F^\#$ to the completion of $\sigma(\mathcal{A})$. In contrast with the scalar case, we cannot however guarantee that $F^\#$ extends to $\sigma(\mathcal{A})$. Nevertheless, we happen to show that if F is the vector measure corresponding to a random measure, this extension is always possible.

(2) Let (Y, \mathcal{B}, ν) be an internal measure space and let $(Y, \mathcal{B}_L, \nu_L)$ stand for the corresponding Loeb measure space. Denote by $L_0(\nu)$ the space of \mathcal{B} -measurable functions from Y to ${}^*\overline{\mathbb{R}}$. As usual, we identify functions that equal ν -almost everywhere.

We consider the external subspaces V_1 and V_2 of $L_0(\nu)$ consisting of ν_L -almost everywhere limited functions and ν_L -almost everywhere infinitesimal functions, respectively; i.e., $f \in V_1$ ($f \in V_2$) whenever there is $U \in \mathcal{B}_L$ such that $\nu_L(Y - U) = 0$ and $f(y)$ is limited (infinitesimal) for all $y \in U$. This definition is sound since if $f \in V_1$ ($f \in V_2$) and $g(y) = f(y)$ for ν -almost all y then $g(y) = f(y)$, implying that g also belongs to V_1 (respectively, to V_2).

The quotient space V_1/V_2 is the *nonstandard hull* of $L_0(\nu)^\#$.

We may identify the nonstandard hull $L_0(\nu)^\#$ with the space $L_0(\nu_L)$ of \mathcal{B}_L -measurable functions from Y to $\overline{\mathbb{R}}$. Namely, we map a coset $f + V_2 \in V_1/V_2$ to ${}^\circ f$, Theorem 6.3.6 implies that this is a linear isomorphism. Consequently,

$$L_0(\nu)^\# = L_0(\nu_L).$$

(3) Assume as before that λ is an internal random measure with respect to (X, \mathcal{A}) and (Y, \mathcal{B}, ν) . Assume further that $\lambda(A, y)$ is limited for all $A \in \mathcal{A}$ and almost all $y \in Y$. Let λ^L be an extension of λ as prompt by Theorem 6.6.10. Note that these two random measures can be considered as the vector measures $\Lambda : \mathcal{A} \rightarrow L_0(\nu)$ and $\Lambda^L : \sigma(\mathcal{A}) \rightarrow L_0(\nu_L)$ acting by the rules $\Lambda(A) := \lambda_A$ and $\Lambda^L(A) := \lambda_A^L$.

6.6.13. Theorem. *The measure $\Lambda^\#$ (see 6.6.12 (1)) is defined soundly and admits an extension to $\sigma(\mathcal{A})$; i.e., there is a vector Loeb measure for Λ . Furthermore, such an extension of $\Lambda^\#$ agrees with Λ^L on $\sigma(\mathcal{A})$.*

◁ Since $\lambda(A, y)$ is limited for almost all $y \in Y$, it follows that the function $\Lambda(A) := \lambda_A$ belongs to V_1 for all $A \in \mathcal{A}$. Hence, $\Lambda^\#$ is defined on \mathcal{A} . Keeping in mind the identification of 6.6.12 (2), for all $A \in \mathcal{A}$ we infer that

$$\Lambda^\#(A) = \Lambda(A)^\# = (\lambda_A)^\# = {}^\circ \lambda_A = \lambda_A^L = \Lambda^L(A),$$

with equality holding almost everywhere. Therefore, $\Lambda^\#$ coincides with Λ^L on \mathcal{A} . Since the measure Λ^L is defined on $\sigma(\mathcal{A})$, we conclude that Λ^L is an extension of $\Lambda^\#$ to $\sigma(\mathcal{A})$. ▷

6.6.14. COMMENTS.

(1) The pseudointegral operators were introduced by Arveson [20] for studying operator algebras on L^2 . They were further studied by Fakhoury [115] (in L^1) and Kalton [208] (in L^p for $0 < p \leq 1$). Various aspects of pseudointegral operators are reflected in [208–211, 441–443, 455–459, 522–524]. Preliminaries to pseudointegral operators are collected in [268].

(2) The main results of this section belong to Troitskiĭ [490].

Chapter 7

Infinitesimals in Harmonic Analysis

In this chapter we elaborate on the technique of hyperapproximation of the Fourier transform on a locally compact abelian group.

We start with the Fourier transform on the reals

$$\mathcal{F} : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R}).$$

In this case the matrix of the discrete Fourier transform is applied to the table of a function f on a grid with knots $-L\Delta_1, \dots, L\Delta_1$. This vector is compared with the table of the Fourier transform $\mathcal{F}(f)$ of F on a grid with knots $-L\Delta, \dots, L\Delta$. We seek the conditions under which the norm of the difference of these two vectors vanishes as Δ and Δ_1 tend to zero and $L\Delta$ and $L\Delta_1$ tend to infinity (or, in other words, this difference is infinitesimal whenever Δ and Δ_1 are infinitesimals and $L\Delta$ and $L\Delta_1$ are infinites). The answer depends essentially on interplay between L , Δ , and Δ_1 .

In fact, the results of 7.1 for the Fourier transform on the reals have a group-theoretic nature, enabling us to abstract them to separable locally compact abelian groups. Hyperapproximation of a locally compact abelian group G is then applied to discrete approximation of the Hilbert space of square-integrable functions on G .

Finally, we apply all these constructions to hyperapproximation of an operator by hyperapproximation of its symbol. We also state results about the limit behavior of Hilbert–Schmidt operators and Schrödinger-type operators.

7.1. Hyperapproximation of the Fourier Transform on the Reals

In this section we consider the possibility of approximating the Fourier transform $\mathcal{F} : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$ by a discrete Fourier transform.

7.1.1. As regards this section, we agree that $X := \{k \in {}^*\mathbb{Z} : -L \leq k \leq L\}$ and $N := 2L + 1$, with L an infinite hypernatural (i.e., $L \approx +\infty$); Δ and Δ' are strictly positive infinitesimals such that $N\Delta \approx +\infty$ and $N\Delta' \approx +\infty$; the functions $j : X \rightarrow {}^*\mathbb{R}$ and $j' : X \rightarrow {}^*\mathbb{R}$ are defined by the rules $j(k) := k\Delta$ and $j'(k) := k\Delta'$.

Thus, (X, j, Δ) and (X, j', Δ') are hyperapproximants to the measure space (\mathbb{R}, Ω, dx) , where Ω is the σ -algebra of Lebesgue-measurable sets, and dx is Lebesgue measure on \mathbb{R} (cf. 6.4.9 (1)).

Consider the internal hyperfinite-dimensional space \mathbb{C}^X and put $E_k(m) := \delta_{km}$ for all $k, m \in X$, so distinguishing the basis $\{E_k : k \in X\}$ for \mathbb{C}^X . Each endomorphism A of \mathbb{C}^X will be given by the matrix of A relative to this basis.

We equip \mathbb{C}^X with the two inner products $(\cdot, \cdot)_\Delta$ and $(\cdot, \cdot)_{\Delta'}$ that are determined from the conditions $(E_k, E_m)_\Delta := \Delta\delta_{km}$ and $(E_k, E_m)_{\Delta'} := \Delta'\delta_{km}$. We also let $\mathcal{L}_{2\Delta}$ stand for \mathbb{C}^X equipped with the inner product $(\cdot, \cdot)_\Delta$ while denoting \mathbb{C}^X with the inner product $(\cdot, \cdot)_{\Delta'}$ by $\mathcal{L}_{2\Delta'}$.

The operator $\Phi : \mathbb{C}^X \rightarrow \mathbb{C}^X$ with matrix $(\exp(-2\pi i\alpha\beta/N))_{\alpha, \beta=-L}^L$ is a *discrete Fourier transform*. In the sequel we will consider the discrete Fourier transform $\Phi_\Delta := \Delta\Phi : \mathcal{L}_{2\Delta} \rightarrow \mathcal{L}_{2\Delta'}$. It is easy to check that

$$(\Phi_\Delta(F), \Phi_\Delta(G))_{\Delta'} = N\Delta\Delta'(F, G)_\Delta.$$

Thus, $\|\Phi_\Delta\| = N\Delta\Delta'$, and so $\Phi_\Delta^\# : \mathcal{L}_{2\Delta}^\# \rightarrow \mathcal{L}_{2\Delta'}^\#$ is a nonzero operator on condition that $0 < N\Delta\Delta' < +\infty$.

Unless otherwise stated, it is assumed in this section that $N\Delta\Delta' \approx 1$. The case of equality holding is especially distinguished by using the notation $\widehat{\Delta} := (N\Delta)^{-1}$ for Δ' . The triple $(X, \widehat{j}, \widehat{\Delta})$ is a hyperapproximant to the measure space (\mathbb{R}, Ω, dx) in this case.

Put $M := j^{-1}(\text{nst}({}^*\mathbb{R}))$ and $\varphi := \text{st} \circ j|_M : M \rightarrow \mathbb{R}$. Then φ induces a mapping $j_2 : L_2(\mathbb{R}) \rightarrow L_2(\nu_\Delta^M) \subset \mathcal{L}_{2, \Delta}^\#$. Recall that if $f \in L_2(\mathbb{R})$ then $j_2(f) = F^\#$, where $F \in \mathcal{S}_2(M)$ is a lifting of $f \circ \varphi$ also called a lifting of f for brevity. Furthermore, if f is bounded and almost everywhere continuous with $|f|^2$ satisfying 6.4.11 (2) then we may take as F the vector $X_\Delta(f) \in \mathbb{C}^X$, with entries $X_\Delta(f)_k := {}^*f(k\Delta)$ for $k \in X$ (cf. 6.4.11).

The entities M', φ', j'_2 , and $X_{\Delta'}(f)$ (as well as $\widehat{M}, \widehat{\varphi}, \widehat{j}_2$, and $X_{\widehat{\Delta}}(f)$) appear likewise on substituting Δ' (respectively, $\widehat{\Delta}$) for Δ .

Finally, we let $\mathcal{F} : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$ stand for the Fourier transform on the reals:

$$\mathcal{F}(f)(y) := \int_{-\infty}^{\infty} f(x) \exp(-2\pi ixy) dx.$$

7.1.2. Theorem. *If $N\Delta\Delta' \approx 1$ then the discrete Fourier transform $\Phi_\Delta : \mathcal{L}_{2, \Delta} \rightarrow \mathcal{L}_{2, \Delta'}$ is a hyperapproximant to $\mathcal{F} : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$.*

◁ Let \mathfrak{M} stand for the set of the characteristic functions of all intervals $[0, a]$ and $[-a, 0]$, with $a > 0$. Then the linear span of \mathfrak{M} is dense in $L_2(\mathbb{R})$ and so we may appeal to 6.5.1 (2).

Put $f := \chi_{[0,a]}$ and let T be an infinite hypernatural satisfying

$$(T - 1)\Delta \leq a < T\Delta.$$

In view of the definition of hyperapproximant, we are left with proving that

$$\Delta' \sum_{k=-L}^L \left| \Delta \sum_{n=0}^{T-1} \exp(-2\pi ink/N) - \int_0^a \exp(-2\pi i x k \Delta') dx \right|^2 \approx 0.$$

It suffices clearly to consider only the terms under the sign $\sum_{k=1}^L$. Moreover, we may replace \int_0^a with $\int_0^{T\Delta}$ since $N\Delta\Delta' \approx 1$ by hypothesis. Therefore,

$$\Delta' \sum_{k=-L}^L \left| \int_a^{T\Delta} \exp(-2\pi i x k \Delta') dx \right|^2 \leq N\Delta^2\Delta' \approx \Delta \approx 0.$$

Simple calculation shows that we are left with proving that

$$\begin{aligned} \Delta' \sum_{k=-L}^L & |\Delta(1 - \exp(-2\pi ikT/N))(1 - \exp(-2\pi ik/N))^{-1} \\ & - (2\pi ik\Delta')^{-1}(1 - \exp(-2\pi ikT\Delta\Delta'))|^2 \approx 0. \end{aligned}$$

Replacing Δ' with $\widehat{\Delta} := (N\Delta)^{-1}$, we obtain

$$\frac{\Delta}{N} \sum_{k=1}^L |1 - \exp(-2\pi ikT/N)|^2 |(1 - \exp(-2\pi ik/N))^{-1} - N/(2\pi ik)|^2 \approx 0.$$

The last formula is obviously valid. Indeed, $0 < 2\pi k/N < \pi$ for all $k \in [1, L]$. Hence, the function $(1 - \exp(-i\varphi))^{-1} - (i\varphi)^{-1}$ has a finite limit as $\varphi \rightarrow 0$, and so f is bounded on $[0, \pi]$. This implies that the validity of the formula we discuss follows from the approximate equality

$$\begin{aligned} \Delta' \sum_{k=1}^L & |(2\pi ik\Delta')^{-1}(1 - \exp(-2\pi ikT\Delta\Delta')) \\ & - N\Delta(2\pi ik)^{-1}(1 - \exp(-2\pi ikT/N))|^2 \approx 0, \end{aligned}$$

which in turn is a consequence of the following two formulas:

$$\frac{1}{\Delta'} \sum_{k=1}^L k^{-2} |(N\Delta\Delta' - 1)(1 - \exp(2\pi ikT/N))|^2 \approx 0$$

and

$$\frac{1}{\Delta'} \sum_{k=1}^L k^{-2} |\exp(-2\pi ikT\Delta\Delta') - \exp(-2\pi ikT/N)|^2 \approx 0.$$

We will prove the first of them, since the second is proved similarly.

Put $\alpha := N\Delta\Delta' - 1 \approx 0$ and $\bar{a} := T\Delta \approx a$. Then the formula under test amounts to the following

$$\frac{\alpha^2}{\Delta'} \sum_{k=1}^L \frac{1}{k^2} \sin^2 \frac{\pi k \bar{a}}{N\Delta} \approx 0.$$

If $\alpha^2 M\Delta' \approx 0$ then the last formula holds since \bar{a} is limited and $\sin^2 x \leq x^2$. If $\alpha^2 M\Delta' \not\approx 0$ then we put $S := \lceil \frac{1}{\alpha\Delta'} \rceil$. In this case $S\Delta'$ is limited and $\alpha^2 S\Delta' \approx 0$. Hence, $S < M$.

We now infer that

$$\frac{\alpha^2}{\Delta'} \sum_{k=1}^L \frac{1}{k^2} \sin^2 \frac{\pi k \bar{a}}{N\Delta} = \frac{\alpha^2}{\Delta'} \sum_{k=1}^S \frac{1}{k^2} \sin^2 \frac{\pi k \bar{a}}{N\Delta} + \frac{\alpha^2}{\Delta'} \sum_{k=S+1}^L \frac{1}{k^2} \sin^2 \frac{\pi k \bar{a}}{N\Delta}.$$

The first term on the right side is infinitesimal because $\alpha^2 S\Delta' \approx 0$, and the second term is infinitesimal too since

$$\frac{\alpha^2}{\Delta'} \sum_{k=S+1}^L \frac{1}{k^2} \leq \frac{\alpha^2}{\Delta'} (S^{-1} - M^{-1})$$

while $S\Delta'$ and $M\Delta'$ are unlimited. \triangleright

7.1.3. We now consider two corollaries of Theorem 7.1.2.

(1) If $N\Delta\Delta' \approx 2\pi h$, where $h > 0$ is a standard real then the discrete Fourier transform $\Phi_\Delta : \mathcal{L}_{2,\Delta} \rightarrow \mathcal{L}_{2,\Delta'}$ is a hyperapproximant to the Fourier transform $\mathcal{F}_h : L_2(\mathbb{R}) \rightarrow L_2(\mathbb{R})$ acting by the rule

$$\mathcal{F}_h(f)(y) := \int_{-\infty}^{\infty} f(x) \exp(-ixy/h) dx.$$

\triangleleft This is immediate by passing from $f \in L_2(\mathbb{R})$ to the function φ defined as $\varphi(t) := f(2\pi ht)$ and replacing Δ with $\Delta/2\pi h$. \triangleright

(2) Suppose that f and $\mathcal{F}(f)$ are bounded almost everywhere continuous functions, with $|f|^2$ and $|\mathcal{F}(f)|^2$ satisfying the condition

$$\lim_{\substack{\Delta \rightarrow 0 \\ A \rightarrow \infty}} \Delta \sum_{|k| > \frac{A}{\Delta}} |f(k\Delta)| = 0.$$

Then

$$\lim_{n \rightarrow \infty} \Delta'_n \sum_{k=-n}^n \left| \int_{-\infty}^{\infty} f(x) \exp(-2\pi i x k \Delta'_n) dx - \Delta_n \sum_{m=-n}^n f(m\Delta_n) \exp(-2\pi i k m/n) \right|^2 = 0$$

for all sequences (Δ_n) and (Δ'_n) such that $\Delta_n \rightarrow 0$, $\Delta'_n \rightarrow 0$, and $n\Delta_n\Delta'_n \rightarrow 1/2$ as $n \rightarrow \infty$.

< This is a standard version of Theorem 7.1.2 immediate from 6.5.10 (3) (also see 6.5.11 (5)). >

(3) We now compare the results on hyperapproximation of the Fourier transform with the results on hyperapproximation of Hilbert–Schmidt operators (cf. 6.5).

The Fourier transform \mathcal{F} is an integral operator with kernel $k_{\mathcal{F}}(x, y) := \exp(-2\pi i xy)$ a bounded analytic function.

If we consider a Hilbert–Schmidt operator A whose kernel $k \in L_2(\mathbb{R}^2)$ is a bounded almost everywhere continuous function with $|k|^2$ satisfying the condition of 6.5.5 (1) (i.e., $k \in \widetilde{\mathcal{H}}_2$) then the operator $A_k : \mathcal{L}_{2,\Delta} \rightarrow \mathcal{L}_{2,\Delta}$ with matrix $(\Delta^* k(\alpha\Delta, \beta\Delta))_{\alpha,\beta=-L}^L$ is a hyperapproximant to A with kernel k .

The discrete Fourier transform Φ_{Δ} , regarded as an operator from $\mathcal{L}_{2,\Delta}$ to $\mathcal{L}_{2,\Delta}$, becomes a hyperapproximant to \mathcal{F} provided that $N\Delta^2 = 1$ (or $N\Delta^2 \approx 1$); cf. Theorem 7.1.2. In this event the matrix of Φ_{Δ} takes the shape

$$\mathcal{A}_k := (\Delta k_{\mathcal{F}}(\alpha\Delta, \beta\Delta))_{\alpha,\beta=-L}^L.$$

The next proposition shows that if we refuse to maintain the above relations between N and Δ then the matrix \mathcal{A}_k may cease determining a hyperapproximant to \mathcal{F} .

7.1.4. If $N\Delta^2 = 2$ then the operator $B : \mathcal{L}_{2,\Delta} \rightarrow \mathcal{L}_{2,\Delta}$ with matrix \mathcal{A}_k is not a hyperapproximant to the Fourier transform \mathcal{F} .

◁ Rewrite the matrix of B as

$$\left(\Delta \exp(-4\pi im/N)\right)_{n,m=-L}^L.$$

Put $f := \chi_{[0;\sqrt{3/2}]}$ and show that

$$\circ(\|B(X_\Delta(f)) - X_\Delta(\mathcal{F}(f))\|_\Delta) > 0.$$

Choose an infinite hypernatural T so that $(T-1)\Delta \leq \sqrt{3/2} < T\Delta$. Easy calculations show that the claim will follow from the inequality

$$\circ\left(\Delta^3 \sum_{m=L-T}^L |1 - \exp(-4\pi imT/N)|^2 |(1 - \exp(-4\pi im/N))^{-1} - N/(4\pi im)|^2\right) > 0.$$

Since $T/N \approx 0$, it is easy that the sum under the standard part operation is at least

$$\left(\Delta^3 \sum_{m=L-T}^L \sin^2 \frac{2\pi mT}{N}\right) \left(\sin^2 \frac{2\pi m}{N}\right)^{-1}.$$

Check now that the standard part of the last hyperreal is strictly positive. To this end, note that $\sin^2(2\pi m/N)$ decreases in m for $L-T \leq m \leq L$. Put $S := [2T/3]$. Then there are some infinitesimals γ and δ such that $\pi T - 3\pi/4 - \gamma \leq 2\pi mT/N \leq \pi T - \pi/2 - \delta$ for all $M-T \leq m \leq M-S$ implying that $\sin^2(2\pi mT/N)$ increases in this interval. Therefore, all terms increase with the so-distinguished indices. Further, the term with index $m := L-T$ is at least $\mathcal{D}\Delta^{-2}$ for some standard real $\mathcal{D} > 0$. It is now obvious that the whole sum is at least $\mathcal{D}(T-S)\Delta$ which is not infinitesimal. ▷

7.1.5. Theorem 7.1.2 will be now abstracted to some class of tempered distributions. In this subsection we put $N := 2L+1$ and let Δ satisfy the additional condition $0 < \circ(N\Delta^2) < +\infty$. If $0 < \circ(N\Delta\Delta') < +\infty$ then this condition also satisfies for N and Δ' .

We start with studying how to assign some element of ${}^*\mathbb{C}^X$ with a given distribution. To this end, consider the operator $\mathcal{D}_d : \mathcal{L}_{2,\Delta} \rightarrow \mathcal{L}_{2,\Delta}$ with matrix $((2\Delta)^{-1}d_{nk})_{n,k=-L}^L$ having the entries

$$d_{nk} := \begin{cases} 1, & \text{if } k = n + 1, \\ -1, & \text{if } k = n - 1, \\ 0 & \text{otherwise,} \end{cases}$$

where the symbols $\dot{+}$ and $\dot{-}$ stand for addition and subtraction in the additive group of the ring ${}^*\mathbb{Z}/N^*\mathbb{Z}$ with the underlying set $\{-L, \dots, L\}$. In other words, if $G := \mathcal{D}_d F$ then $G(k) = (F(k + 1) - F(k - 1))(2\Delta)^{-1}$. Clearly, $\|\mathcal{D}_d\| \approx +\infty$, and so it is impossible to speak of $\mathcal{D}_d^\#$. To proceed further, we prove a few auxiliary propositions.

(1) If $G^{(n)} = \mathcal{D}_d^n F$ then

$$G^{(n)}(k) = \frac{1}{(2\Delta)^n} \sum_{r=0}^n (-1)^r \binom{n}{r} F(k \dot{+} n \dot{-} 2r).$$

Furthermore, if $|k| \leq L - n$, then symbols $\dot{+}$ and $\dot{-}$ may be replaced with $+$ and $-$, respectively.

(2) If

$$\Delta^{-2} F(\pm(L - t)) \approx 0$$

for all standard s and t then $G^{(n)}(\pm(L - m)) \approx 0$ for all standard n and m .

(3) If $f \in S(\mathbb{R})$ then $\Delta^{-s} f(\pm(L - t)\Delta) \approx 0$ for all standard s and t .

\triangleleft Since $L\Delta \approx +\infty$ and t is standard; therefore, $(L - t)\Delta \approx +\infty$. Considering that $f \in S(\mathbb{R})$, obtain $[(L - t)\Delta]^{-s} f(\pm(L - t)\Delta) \approx 0$. Thus, $0 \approx \Delta^{-s} [(L - t)\Delta^2]^s f(\pm(L - t)\Delta)$. We are done on using the condition $0 < \circ(N\Delta^2) < +\infty$. \triangleright

(4) If f belongs to the Schwartz space $S(\mathbb{R})$ and n is a standard natural then

$$\|\mathcal{D}_d^n X_\Delta(f) - X_\Delta(f^{(n)})\|_{2,\Delta} \approx 0.$$

\triangleleft By (2) and (3),

$$\Delta \sum_{|k| > L - n} |(\mathcal{D}_d^n X_\Delta(f))(k) - X_\Delta(f^{(n)})(k)|^2 \approx 0.$$

In view of (1) it suffices to show that

$$0 \approx S = \Delta \sum_{k=-L+n}^{L-n} \left| \frac{1}{(2\Delta)^n} \sum_{r=0}^n (-1)^r \binom{n}{r} f((k + n - 2r)\Delta) - f^{(n)}(k\Delta) \right|^2.$$

By the Taylor Theorem,

$$\begin{aligned} f(k\Delta + (n - 2r)\Delta) &= \sum_{s=0}^n \frac{f^{(s)}(k\Delta)}{s!} (n - 2r)^s \Delta^s \\ &\quad + \frac{f^{(n+1)}(\xi_r)}{(n + 1)!} (n - 2r)^{n+1} \Delta^{n+1}. \end{aligned}$$

Hence,

$$\begin{aligned} & \frac{1}{(2\Delta)^n} \sum_{r=0}^n (-1)^r \binom{n}{r} f((k+n-2r)\Delta) \\ &= \frac{1}{(2\Delta)^n} \sum_{s=0}^n \frac{f^{(s)}(k\Delta)}{s!} \Delta^s \sum_{r=0}^n (-1)^r \binom{n}{r} (n-2r)^s \\ & \quad + \frac{\Delta}{2^n} \sum_{r=0}^n \frac{f^{(n+1)}(\xi r)}{(n+1)!} (n-2r)^{n+1}. \end{aligned}$$

Using the easy formula

$$\sum_{r=0}^n (-1)^r r^s \binom{n}{r} = \begin{cases} 0, & \text{if } s < n, \\ (-1)^n n!, & \text{if } s = n, \end{cases}$$

infer that the first term on the right side of the penultimate formula is equal to $f^{(n)}(k\Delta)$. Since $f^{(n+1)}$ is bounded on \mathbb{R} , the modulus of the second term on the right side is bounded above by $B\Delta$ with B a standard real. Consequently, $S \leq \Delta \cdot 2(L-n)B^2\Delta^2$ and we are done on appealing to (1). \triangleright

7.1.6. We now define the sequence $(\mathcal{L}^{(n)})_{n \in \mathbb{N}}$ of external subspaces of $\mathcal{L}_{2,\Delta}$ as follows:

$$\begin{aligned} \mathcal{L}^{(0)} &:= \left\{ F \in \mathcal{L}_{2,\Delta} : (\forall^{\text{st}} a)(\exists^{\text{st}} C) \left(\Delta \sum_{k=-[a/\Delta]}^{[a/\Delta]} |F(k)|^2 < C \right) \right\}; \\ \mathcal{L}^{(n+1)} &:= \mathcal{D}_d(\mathcal{L}^{(n)}); \\ \mathcal{L}^{(\sigma)} &:= \bigoplus_{n=0}^{\infty} \mathcal{L}^{(n)}. \end{aligned}$$

Assume that $F \in {}^*\mathbb{C}^X$ is such that

$$\circ \left(\Delta \sum_{k=-L}^L F(k) {}^* f(k\Delta) \right) = \circ(F, \overline{X_\Delta(f)}) < +\infty$$

for every standard $f \in C_0^\infty(\mathbb{R})$.

The formula

$$\psi_F^\Delta(f) := \circ(F, X_\Delta(f))_\Delta$$

defines a (possibly discontinuous) linear functional $\psi_F^\Delta : C_0^\infty(\mathbb{R}) \rightarrow \mathbb{C}$.

Theorem. *The following hold:*

- (1) $\psi_F^\Delta \in (C_0^\infty(\mathbb{R}))'$ for all $F \in \mathcal{L}^{(\sigma)}$;
- (2) If $f \in (C_0^\infty(\mathbb{R}))'$ and $f = \varphi^{(k)}$ for some regular distribution φ ($k \geq 0$) then there is some F in $\mathcal{L}^{(\sigma)}$ satisfying $\psi_F^\Delta = f$;
- (3) $\psi_{\mathcal{D}_d F}^\Delta = (\psi_F^\Delta)'$ for all $F \in \mathcal{L}^{(\sigma)}$.

◁ Given $A \leq L$, put $(F, G)_\Delta^A := \Delta \sum_{k=-A}^A F(k) \overline{G(k)}$. Suppose that $f, f_n \in C_0^\infty(\mathbb{R})$ and $f_n \rightarrow f$ in $C_0^\infty(\mathbb{R})$ as $n \rightarrow \infty$. Then there is a standard real $a > 0$ such that the supports $\text{supp}(f_n)$ and $\text{supp}(f)$ lie in $[-a, a]$ and (f_n) converges uniformly to f . Since the supports of f_n and f are compact; therefore, $\|f_n\|_{L_2}$ and $\|f\|_{L_2}$ are limited and $\|f - f_n\|_{L_2} \rightarrow 0$ as $n \rightarrow \infty$.

Put $A := [a/2]$. Then

$$\|f\|_{L_2} = {}^\circ\|X_\Delta(f)\|_{2,\Delta}^A, \quad \|f_n\|_{L_2} = {}^\circ\|X_\Delta(f_n)\|_{2,\Delta}^A, \\ {}^\circ\|X_\Delta(f - f_n)\|_{2,\Delta}^A \xrightarrow{n \rightarrow \infty} 0.$$

Now, if $F \in \mathcal{L}^{(0)}$, then

$$|\psi_F^\Delta(f)| = |{}^\circ(F, \overline{X_\Delta(f)})_\Delta| = |{}^\circ(F, \overline{X_\Delta(f)})_\Delta^A| \leq {}^\circ\|F\|_{2,\Delta}^A {}^\circ\|X_\Delta(f)\|_{2,\Delta}^A.$$

From the definition of $\mathcal{L}^{(0)}$ it follows that ${}^\circ\|F\|_{2,\Delta}^A$ is limited as well ${}^\circ\|X_\Delta(f)\|_{2,\Delta}^A$ (we have seen this earlier). Hence, $\psi_F^\Delta(f)$ is limited for all $f \in C_0^\infty(\mathbb{R})$. By analogy, $\psi_F^\Delta(f_n - f) \rightarrow 0$ as $n \rightarrow \infty$, and so $\psi_F^\Delta \in (C_0^\infty(\mathbb{R}))'$.

Note further that $(\mathcal{D}_d^n F, \overline{X_\Delta(f)}) = (-1)^n (F, \overline{\mathcal{D}_d^n X_\Delta(f)})$. From 7.1.5 (4) it follows that $\mathcal{D}_d^n X_\Delta(f) = X_\Delta(f^{(n)}) + T$, where $\|T\|_{2,\Delta} \approx 0$. Moreover, since $\text{supp } f^{(n)} \subset [-a, a]$; therefore, $X_\Delta(f^{(n)})(k) = 0$ for $|k| > A$.

As follows from 7.1.5 (1), $\mathcal{D}_d^n X_\Delta(f)(k) = 0$ for $|k| > A + n$. Since n is standard, infer that $T(k) = 0$ for $|k| > [b/\Delta]$ and each standard $b > a$. Thus, if $B = [b/\Delta]$ then $(F, T)_\Delta = (F, T)_\Delta^B \approx 0$, because $\|F\|_{2,\Delta}^B$ is limited and $\|T\|_{2,\Delta}^B = \|T\|_{2,\Delta} \approx 0$. Hence, $(\mathcal{D}_d^n F, \overline{X_\Delta(f)}) \approx (-1)^n (F, \overline{X_\Delta(f^{(n)})})$, implying that $\psi_{\mathcal{D}_d^n F}^\Delta(f) = \psi_F^{\Delta(n)}(f)$. This yields (1) and (3).

To prove (2) it suffices to observe that there is no loss of generality in assuming φ continuous, implying that $X_\Delta(\varphi) \in \mathcal{L}^{(0)}$. ▷

7.1.7. Put $\mathfrak{E}^{(0)} := \text{ltd}(\mathcal{L}_{2,\Delta})$; $\mathfrak{E}^{(n+1)} := \mathcal{D}_d \mathfrak{E}^{(n)}$, and

$$\mathfrak{E}^{(\sigma)} := \bigoplus_{n=0}^{\infty} \mathfrak{E}^{(n)}.$$

Obviously, $\mathfrak{E}^{(0)} \subset \mathcal{L}^{(0)}$, and so $\mathfrak{E}^{(n)} \subset \mathcal{L}^{(n)}$ for all n . Therefore, $\mathfrak{E}^{(\sigma)} \subset \mathcal{L}^{(\sigma)}$. The next two propositions may be proven along the lines of the proof of the above theorem.

(1) If $F \in \mathfrak{C}^{(\sigma)}$ then ψ_F^Δ is a tempered distribution.

(2) If $f = \varphi^{(k)}$ and $\varphi \in L_2(\mathbb{R})$ then there is an element F in $\mathfrak{C}^{(\sigma)}$ satisfying $f = \psi_F^\Delta$.

Supposing that Δ' enjoys the condition $N\Delta\Delta' \approx 1$, we again consider the discrete Fourier transform $\Phi_\Delta : \mathcal{L}_{2,\Delta} \rightarrow \mathcal{L}_{2,\Delta'}$. Since Φ_Δ has a limited norm; therefore, $\Phi_\Delta(\mathfrak{C}^{(0)}) = \text{ltd}(\mathcal{L}_{2,\Delta'}) = \widehat{\mathfrak{C}}^{(0)}$ and $\Phi_\Delta^{-1}(\widehat{\mathfrak{C}}^{(0)}) = \mathfrak{C}^{(0)}$.

Put $\mathcal{M}_d := \Phi_\Delta \mathcal{D}_d \Phi_\Delta^{-1} : \mathcal{L}_{2,\Delta'} \rightarrow \mathcal{L}_{2,\Delta'}$ and define the sequence $(\widehat{\mathfrak{C}}^{(n)})_{n \in \mathbb{N}}$ of external subspaces of $\mathcal{L}_{2,\Delta'}$ by letting $\widehat{\mathfrak{C}}^{(n+1)} := \mathcal{M}_d(\widehat{\mathfrak{C}}^{(n)})$. It is now obvious that $\Phi_\Delta(\mathfrak{C}^{(n)}) = \widehat{\mathfrak{C}}^{(n)}$ and $\widehat{\mathfrak{C}}^{(\sigma)} = \bigoplus_{n=0}^{\infty} \widehat{\mathfrak{C}}^{(n)} = \Phi_\Delta \mathfrak{C}^{(\sigma)}$. Straightforward calculation shows that the operator \mathcal{M}_d may be given by the matrix

$$\left(\frac{i}{\Delta} \sin \frac{2\pi\beta}{N} \delta_{\alpha\beta} \right)_{\alpha,\beta=-L}^L.$$

Given $G \in \mathbb{C}^X$ and $f \in C_0^\infty(\mathbb{R})$, we define $\psi_G^{\Delta'}(f)$ as in 7.1.6 on replacing Δ with Δ' and F with G .

7.1.8. If $f \in S(\mathbb{R})$ then

$$\circ \|\mathcal{M}_d^n(X_{\Delta'}(f)) - X_{\Delta'}(\mathcal{M}^n(f))\| = 0.$$

◁ Considering the formulas of 7.1.7 for the matrix of \mathcal{M}_d , we are left with demonstrating only that

$$\Delta' \sum_{\beta=-L}^L \left| \left(\frac{1}{\Delta} \sin \frac{2\pi\beta}{N} \right)^n f(\beta\Delta') - (2\pi\beta\Delta')^n f(\beta\Delta') \right|^2 \approx 0.$$

Show first that if $T < L$ but $T\Delta' \approx +\infty$ then

$$\mathscr{W} := \Delta' \sum_{|\beta|>T} |f(\beta\Delta')|^2 \frac{1}{\Delta^{2n}} \left| \left(\sin \frac{2\pi\beta}{N} \right)^n - (2\pi\beta\Delta\Delta')^n \right|^2 \approx 0.$$

Indeed, since $N\Delta\Delta' \approx 1$, there is some standard C satisfying

$$\begin{aligned} & \left| \left(\sin \frac{2\pi\beta}{N} \right)^n - (2\pi\beta\Delta\Delta')^n \right|^2 \\ &= \left| \frac{\beta}{N} \right|^{2n} \left| \left(\frac{\beta}{N} \right)^{-n} \left(\sin \frac{2\pi\beta}{N} \right)^n - (2\pi N\Delta\Delta')^n \right|^2 \leq C \left| \frac{\beta}{N} \right|^{2n}. \end{aligned}$$

This enables us to estimate \mathscr{W} as follows:

$$\mathscr{W} \leq C_1 \Delta' \sum_{|\beta| > T} |f(\beta \Delta')|^2 |\beta \Delta'|^{2n}.$$

Put $\varphi(x) := x^{2n} f(x)$. Since $f(x) \in S(\mathbb{R})$; therefore, $\varphi(x) \in L_2(\mathbb{R})$. Moreover, φ is a bounded continuous function satisfying the equality under proof. Consequently, the internal function $G : X \rightarrow \mathbb{C}$, given by $G(\beta) := *f(\beta \Delta') (\beta \Delta')^n$, is a lifting of φ , and $G \in \mathcal{S}_2(M)$. This implies that the right side of the above upper estimate for \mathscr{W} is in fact an infinitesimal, yielding $\mathscr{W} \approx 0$.

By 4.6.11 there is some $a \approx +\infty$ such that $N \Delta' \stackrel{a}{\approx} +\infty$ and $N \Delta \Delta' - 1 \stackrel{a}{\approx} 0$. Put $T := [a/\Delta']$. Then T satisfies the preceding conditions, and so to complete the proof it suffices in view of the boundedness of f to show that

$$\mathscr{W}_1 := \Delta' \sum_{\beta=1}^T \left| \left(\frac{1}{\Delta} \sin \frac{2\pi\beta}{N} \right)^n - (2\pi\beta\Delta')^n \right|^2 \approx 0.$$

If $1 \leq \beta \leq T$ then $0 < \beta/N < a/(N\Delta') \stackrel{a}{\approx} 0$, implying that is

$$\left(\sin \frac{2\pi\beta}{N} \right) \left(\frac{2\pi\beta}{N} \right)^{-1} = 1 - \alpha_\beta,$$

where $\alpha_\beta \stackrel{a}{\approx} 0$. Hence, $\Delta^{-n} \left(\sin \frac{2\pi\beta}{N} \right)^n = (1 - \alpha_\beta) (2\pi\beta\Delta')^n (N\Delta\Delta')^{-n}$. By the choice of a there is some $\delta \stackrel{a}{\approx} 0$ satisfying $(N\Delta\Delta')^{-n} = 1 + \delta$. Finally,

$$\Delta^{-n} \left(\sin \frac{2\pi\beta}{N} \right)^n = (1 + \gamma_\beta) (2\pi\beta\Delta')^n,$$

where $\gamma_\beta \stackrel{a}{\approx} 0$. If $\gamma = \max\{|\gamma_\beta| : 1 \leq \beta \leq T\}$ then $\gamma \stackrel{a}{\approx} 0$, yielding

$$\mathscr{W}_1 \leq \Delta' \gamma^2 \sum_{\beta=1}^T (2\pi\beta\Delta')^{2n} \leq \Delta' \gamma^2 (2\pi)^n (T\Delta')^{2n} \leq (2\pi)^n \gamma^2 a^{2n} \Delta' \approx 0,$$

which completes the proof. \triangleright

7.1.9. Theorem. *The following hold:*

- (1) $\psi_G^{\Delta'} \in (S(\mathbb{R}))'$ for all $G \in \widehat{\mathfrak{C}}^{(\sigma)}$;
- (2) $\mathcal{F}(\psi_F^{\Delta'}) = \psi_{\Phi_{\Delta'}(F)}^{\Delta'}$ for all $F \in \mathfrak{C}^{(\sigma)}$.

◁ (1): This is proved in much the same way as the theorem of 7.1.6.

(2): Denote the differentiation operator in $S(\mathbb{R})$ by \mathcal{D} and let $\mathcal{M} := \mathcal{F}\mathcal{D}\mathcal{F}^{-1}$. Then $\mathcal{M}(f)(x) := 2\pi ixf(x)$. It obviously suffices to check (2) for $F := \mathcal{D}_d^n G$, where $G \in \mathfrak{C}^{(\sigma)}$. In this event, everything results however from the next chain of equalities:

$$\begin{aligned}
\psi_{\Phi_{\Delta}(F)}^{\Delta'}(f) &= \circ(\Phi_{\Delta}(\mathcal{D}_d^n G), \overline{X_{\Delta'}(f)}) = \circ(\mathcal{D}_d^n G, \Phi_{\Delta}^{-1}(X_{\Delta'}(\bar{f}))) \\
&= (-1)^{n\circ}(G, \mathcal{D}_d^n \Phi_{\Delta}^{-1} X_{\Delta'}(\bar{f})) \\
&= (-1)^{n\circ}(G, \Phi_{\Delta}^{-1} \mathcal{M}_d X_{\Delta'}(\bar{f})) \\
&= (-1)^{n\circ}(G, \Phi_{\Delta}^{-1} X_{\Delta'}(\mathcal{M}^n(\bar{f}))) \\
&= (-1)^{n\circ}(G, X_{\Delta'}(\mathcal{F}^{-1} \mathcal{M}^n(\bar{f}))) \\
&= (-1)^{n\circ}(G, X_{\Delta}(\mathcal{D}^n \mathcal{F}^{-1}(\bar{f}))) \\
&= (-1)^{n\circ}(G, \mathcal{D}_d^n X_{\Delta}(\mathcal{F}^{-1}(\bar{f}))) \\
&= \circ(\mathcal{D}_d^n G, X_{\Delta}(\mathcal{F}^{-1}(\bar{f}))) = \circ(F, X_{\Delta}(\mathcal{F}^{-1}(\bar{f}))) \\
&= \psi_F^{\Delta}(\mathcal{F}^{-1}(\bar{f}))(\mathcal{F}\psi_F^{\Delta}(f)).
\end{aligned}$$

The proof is complete. ▷

7.1.10. COMMENTS.

(1) The condition $N\Delta\Delta' \approx 1$ of Theorem 7.1.2 arises also in the celebrated Kotel'nikov Theorem asserting that if the spectrum of a bounded function f belongs to the interval $[-a, a]$ then f is completely determined from its values on the set $\{n\lambda : -\infty < n < +\infty\}$, with $\lambda \leq (2a)^{-1}$, by the formula

$$f(t) = \sum_{k=-\infty}^{\infty} f(k\lambda) \frac{\sin 2\pi a(t - k\lambda)}{2\pi a(t - k\lambda)}.$$

In our case, the values of f are calculated at the points $k\Delta$, with $\Delta \approx 1/(N\Delta')$, and it is easy to see that $N\Delta'$ is exactly the length of the interval where we consider $\mathcal{F}(f)$.

(2) The condition $N\Delta\Delta' \approx 2\pi h$ of 7.1.3(1) is closely tied with the *uncertainty principle* of quantum mechanics.

Consider the one-parameter groups of unitary operators $U(u) := \exp(-iuP)$ and $V(v) := \exp(-ivQ)$, where Q and P are the coordinate and momentum operators; i.e., Q is the multiplication by the independent variable, and $P := \frac{h}{i} \frac{d}{dx}$, where $h > 0$ is a distinguished standard real, the *Planck constant*. We recall in this event that $U(u)\varphi(x) = \varphi(x - uh)$, $V(v)\varphi(x) = \exp(-ivx)\varphi(x)$ and the commutation relations

$$U(u)V(v) = \exp(ihuv)V(v)U(u)$$

hold which may be regarded as one of the form of the uncertainty principle.

Define the hyperfinite-rank operators $U_d, V_d : \mathbb{C}^X \rightarrow \mathbb{C}^X$ with $(U_d F)(k) := F(k \dot{-} 1)$ and V_d the diagonal matrix

$$(\exp(-2\pi ink/N)\delta_{nk})_{n,k=-L}^L.$$

It is easy to check that the operators U_d^r and V_d^m satisfy the following commutation relations:

$$U_d^r V_d^m = \exp(2\pi irm/N) V_d^m U_d^r \quad (r, m \in {}^*\mathbb{Z}/N^*\mathbb{Z}).$$

If $\exp(2\pi irm/N) \approx \exp(\pi ihuv)$ then the last relation transforms into the above commutation relation for U and V . However, for the condition $\exp(2\pi irm/N) \approx \exp(\pi ihuv)$ to hold we need relate the quantities $r, m, u,$ and v ; the details are collected in the following proposition.

(3) *If $r, m \in {}^*\mathbb{Z}/N^*\mathbb{Z}$ are such that $r\Delta \approx uh$ and $2\pi m\widehat{\Delta} \approx v$, with u and v standard reals; then U_d^r and $V_d^m : \mathcal{L}_{2,\Delta} \rightarrow \mathcal{L}_{2,\Delta}$ are hyperapproximants to $U(u)$ and $V(v)$, respectively.*

◁ As \mathfrak{M} in 6.5.1 (2) we take the set of the characteristic functions of the closed intervals. By definition it is clear that $U(u)(\chi_{[a,b]}) = \chi_{[a+uh, b+uh]}$. Note that if $k\Delta$ and $m\Delta$ are nearstandard then $|k|, |m|,$ and $|k - m|$ are all less than L . Hence, $k \dot{-} m = k - m$ and $U_d^r(F)(k) = F(k - r)$. If $f = \chi_{[a,b]}$ then

$$X_\Delta(f)(n) = \begin{cases} 0, & \text{if } n\Delta \notin {}^*[a, b], \\ 1, & \text{if } n\Delta \in {}^*[a, b], \end{cases}$$

implying that

$$U_d^r(X_\Delta(f))(n) = \begin{cases} 0, & \text{if } n\Delta \notin {}^*[a + r\Delta, b + r\Delta], \\ 1, & \text{if } n\Delta \in {}^*[a + r\Delta, b + r\Delta]. \end{cases}$$

Similarly,

$$X_\Delta(U(u)(f))(n) = \begin{cases} 0, & \text{if } n\Delta \notin {}^*[a + uh, b + uh], \\ 1, & \text{if } n\Delta \in {}^*[a + uh, b + uh]. \end{cases}$$

It is now obvious that

$$\Delta \sum_{k=-L}^L |U_d^r(X_\Delta(f))(n) - X_\Delta(U(u)(f))(n)|^2 = 2\Delta [|(uh - r\Delta)/\Delta|] \approx 0,$$

since $uh \approx r\Delta$.

Clearly, $V(v) = \mathcal{F}_h^{-1}U^{-1}(v)\mathcal{F}_h$ (cf. 7.1.3 (1)) and $V_d = \Phi_\Delta^{-1}U_d^{-1}\Phi_\Delta$. Put $\Delta_1 := 2\pi p\widehat{\Delta}$ and regard U_d^{-1} as an operator from \mathcal{L}_{2,Δ_1} to \mathcal{L}_{2,Δ_1} . By hypotheses, $m\Delta_1 \approx vh$ and, as was proven, U_d^{-m} is a hyperapproximant to $U^{-1}(v)$. By 7.1.3 (1), the operator $\Phi_\Delta : \mathcal{L}_{2,\Delta} \rightarrow \mathcal{L}_{2,\Delta_1}$ is a hyperapproximant to \mathcal{F}_h ; i.e., $V_d^m := \Phi_\Delta^{-1}U_d^{-m}\Phi_\Delta$ is a hyperapproximant to $\mathcal{F}_h^{-1}U^{-1}(v)\mathcal{F}_h = V(v)$. ▷

(4) Theorem 7.1.9 suggests an approach to approximating the Fourier transform of tempered distributions by the discrete Fourier transform. Unfortunately, this approach is successful only the primitives of the functions in $L_2(\mathbb{R})$. Fortunately, this class may be enriched as shown by the following example.

Assume that $\varphi = 1$. Hence, φ is a regular distribution and by 7.1.8, $\varphi = \psi_F^\Delta$, where $F := X_\Delta(f)$, i.e., $F(k) = 1$ for all $k \in X$. Since $\mathcal{F}(1) = \delta$; therefore, $\Phi_\Delta(F)(k) = N\Delta\delta_{k0}$. Consequently, if $G := \Phi_\Delta(F)$ then $\psi_G^{\Delta'}(f) = {}^\circ(G, X_\Delta(f))_{\Delta'} = {}^\circ(N\Delta\Delta'^*f(0)) = f(0) = \delta(f)$, implying that $\mathcal{F}(\psi_F) = \psi_{\Phi_\Delta(F)}^{\Delta'}$.

7.2. A Nonstandard Hull of a Hyperfinite Group

In this section we study a construction that assign to a hyperfinite group some locally compact group.

7.2.1. We consider the additive group G of the ring ${}^*\mathbb{Z}/N{}^*\mathbb{Z}$ with the underlying set $\{-L, \dots, L\}$; cf. 7.1.1. Clearly, G is an internal hyperfinite abelian group. Distinguish the two natural external subgroups $G_0 := \{k \in G : k\Delta \approx 0\}$ and $G_f := \{k \in G : {}^\circ|k\Delta| < +\infty\}$ of G . It is easy to see that G_0 is the intersection of a countable family of internal sets and G_f , the union of a countable family of internal sets:

$$G_0 = \bigcap_{n \in \mathbb{N}} j^{-1}({}^*(-n^{-1}, n^{-1})), \quad G_f = \bigcup_{n \in \mathbb{N}} j^{-1}({}^*[-n, n]).$$

Further, the mapping $\text{st} : G_f \rightarrow \mathbb{R}$ is an epimorphism, and $\ker(\text{st}) = G_0$, i.e., is $\mathbb{R} \simeq G_f/G_0$.

Suppose that A is an internal set coincident with $j^{-1}({}^*(a, b))$ for some $a, b \in \mathbb{R}$. In this event $\text{st}(A) = [a, b]$. We now define the external set $A^\circ := \{c \in A : c + G_0 \subset A\}$. It is easy to check that $\text{st}(A^\circ) = (a, b)$. A similar definition of A° applies to an arbitrary internal set A . In this event it is an easy matter to that $\text{st}(A)$ is closed and $\text{st}(A^\circ)$ is open. Thus,

$$\{\text{st}(A^\circ) : A \text{ is an internal subset of } G_f\}$$

is a base for the topology of \mathbb{R} . We also see that ${}^\circ(\Delta|j^{-1}({}^*[a, b])) = b - a$. This implies straightforwardly that $(G_f, S_\Delta^{G_f}, \nu_\Delta^{G_f})$ is a σ -finite subspace of the Loeb measure space $(G, S_\Delta, \nu_\Delta)$ (cf. 6.3.11).

In this case $\nu_\Delta^{G_f}$ is an invariant measure on G_f . Moreover, the preceding equality shows that $\text{st} : G_f \rightarrow \mathbb{R}$ is measure-preserving provided that we view Lebesgue measure as the Haar measure on \mathbb{R} .

Let \widehat{G} stand for the character group of G . Then the internal mapping $n \mapsto \chi_n$, with $\chi_n(m) := \exp(2\pi imn/N)$ for $n, m \in G$, is an isomorphism from G to \widehat{G} . This claim follows by transfer since it holds for every standard N .

In order for a character $\chi_n : G \rightarrow {}^*\mathbb{C}$ to induce a character $\varkappa : \mathbb{R} \rightarrow \mathbb{C}$ by composition with the homomorphism st it is necessary and sufficient that $\chi_n|_{G_0} \approx 1$. With this in mind, we naturally distinguish the external subgroup $H_f := \{\chi \in \widehat{G} : \chi|_{G_0} \approx 1\}$ of \widehat{G} and define the monomorphism $\widehat{\text{st}} : H_f \rightarrow \widehat{\mathbb{R}}$ by the rule $\widehat{\text{st}}(\chi)(n) := {}^\circ\chi({}^\circ(n\Delta))$ for $n \in \mathbb{N}$ and $\chi \in \widehat{G}$. Thereby, $\ker(\widehat{\text{st}}) = H_0 := \{\chi \in H_f : \chi|_{G_f} \approx 1\}$ and so $H_f/H_0 \subset \widehat{\mathbb{R}}$.

(1) If $n \in G$ then

$$\begin{aligned} \chi_n \in H_f &\leftrightarrow {}^\circ(|n\widehat{\Delta}|) < +\infty; \\ \chi_n \in H_0 &\leftrightarrow n\widehat{\Delta} \approx 0, \end{aligned}$$

with $\widehat{\Delta} := (N\Delta)^{-1}$.

◁ Indeed, if $n\widehat{\Delta}$ is a limited hyperreal then $\chi_n(m) = \exp(2\pi imn/N) = \exp(2\pi im\Delta n\widehat{\Delta}) \approx 1$, because $m\Delta \approx 0$.

Conversely, suppose that $n\widehat{\Delta}$ is unlimited. Put $m := [N/(2\pi)]$. Clearly, $m\widehat{\Delta} \approx 0$. In this case $\chi_n(m) = \exp 2\pi(\frac{N}{2\pi})\Delta \frac{n}{N\Delta} \approx -1$, because $0 \leq \alpha < 1$, which is a contradiction. The proof of the other equivalence proceeds by analogy. ▷

Thus, the groups \widehat{G}_f and \widehat{G}_0 are constructed just like the groups G_f and G_0 , and so $\widehat{G}_f/\widehat{G}_0 \cong \widehat{\mathbb{R}} \cong \mathbb{R}$.

The Fourier transform $\mathcal{F} : L_2(\mathfrak{G}) \rightarrow L_2(\widehat{\mathfrak{G}})$ for an arbitrary locally compact abelian group \mathfrak{G} is defined as $\mathcal{F}(f)(\chi) := (f, \widehat{\chi})$. Therefore, Theorem 7.1.2 clearly admits a group-theoretic interpretation. We so proceed to the general situation.

Let G be an internal hyperfinite abelian group, and let G_0 and G_f be subgroups of G such that $G_0 \subset G_f$ and the following hold:

- (A) There is a sequence $(A_n)_{n \in \mathbb{N}}$ of internal sets such that $A_n \subset G_f$ for all $n \in \mathbb{N}$ and $G_0 = \bigcap \{A_n : n \in \mathbb{N}\}$;
- (B) There is a sequence $(B_n)_{n \in \mathbb{N}}$ of internal sets such that $B_n \supset G_0$ for all $n \in \mathbb{N}$ and $G_f = \bigcup \{B_n : n \in \mathbb{N}\}$.

Observe that these subgroups G_0 and G_f may be internal as well as external.

(2) If G_0 and G_f are subgroups of G , with $G_0 \subset G_f$, satisfying (A) and (B) then there is a countable sequence $(C_n)_{n \in \mathbb{Z}}$ of symmetric internal subsets of G such that:

- (a) $\bigcap_{n \in \mathbb{Z}} C_n = G_0$;
- (b) $\bigcup_{n \in \mathbb{Z}} C_n = G_f$;
- (c) $C_n + C_n \subset C_{n+1}$ ($n \in \mathbb{Z}$).

Further, if F is an internal subset of G then

- (d) $F \subset G_f \leftrightarrow (\exists n \in \mathbb{Z})(F \subset C_n)$;
- (e) $F \supseteq G_0 \leftrightarrow (\exists n \in \mathbb{Z})(F \supseteq C_n)$.

◁ These auxiliary facts are easy by saturation. ▷

In the sequel we will work with some sequence $(C_n)_{n \in \mathbb{Z}}$ satisfying the conditions of (2). If $F \subset G_f$, then we put $\overset{\circ}{F} := \{g \in G_f : g + G_0 \subset F\}$. The following is immediate from (2).

(3) *If F is an internal subset of G_f then $(\forall g \in G)(g \in \overset{\circ}{F} \leftrightarrow (\exists m \in \mathbb{Z})(g + C_m \subset F))$.*

We denote the family of all internal subsets of G_f by $\text{In}(G_f)$ and put $\text{In}_0(G_f) := \{F \in \text{In}(G_f) : G_0 \subset F\}$. Put $G^\# := G_f/G_0$ and let $j : G_f \rightarrow G^\#$ denote the quotient homomorphism. If $g \in G_f$ and $A \subset G_f$ then instead of $j(g)$ and $j(A)$ we will write $g^\#$ and $A^\#$, respectively.

7.2.2. Theorem. *The following hold:*

- (1) *The family $\mathfrak{U} := \{\overset{\circ}{F}^\# : F \in \text{In}(G_f)\}$ is a neighborhood base of zero for some uniform topology compatible with the group structure on $G^\#$. This topology is called canonical below;*
- (2) *If $F \in \text{In}(G_f)$ then $F^\#$ is closed;*
- (3) *The topological group $G^\#$ is complete.*

◁ (1): From the general theory of topological groups (see, for instance, [402]) it follows that some system \mathfrak{U} of subsets of an abelian group is a neighborhood base of zero whenever the following hold:

- (a) $\bigcap \{U : U \in \mathfrak{U}\} = \{e\}$;
- (b) $(\forall U, V \in \mathfrak{U})(\exists W \in \mathfrak{U})(W \subset U \cap V)$;
- (c) $(\forall U \in \mathfrak{U})(\exists V \in \mathfrak{U})(V - V \subset U)$;
- (d) $(\forall U \in \mathfrak{U})(\forall \xi \in U)(\exists V \in \mathfrak{U})(V + \xi \subset U)$.

We now check this in our environment.

(a): If $g^\# \in \bigcap \{\overset{\circ}{F}^\# : F \in \text{In}_0(G_f)\}$ then $((g + G_0) \cap \overset{\circ}{F} \neq \emptyset)$ for all $F \in \text{In}_0(G_f)$ as follows from the fact that $j^{-1}(g^\#) = g + G_0$ and $j^{-1}(\overset{\circ}{F}^\#) = \overset{\circ}{F} + G_0 = \overset{\circ}{F}$ (see the definition of $\overset{\circ}{F}$). Consequently, there is some $g_1 \in G_0$ satisfying $(g + g_1 + G_0 \subset F)$, and so $g \in F$. Thus, $g \in \bigcap \text{In}_0(G_f) = G_0$.

(b): This ensues from the inclusion $(F_1 \cap F_2)^{\circ\#} \subset \overset{\circ}{F}_1^\# \cap \overset{\circ}{F}_2^\#$, which is checked as above.

(c): If $F \in \text{In}_0(G_f)$ then there is some $n \in \mathbb{Z}$ satisfying $C_n + C_n + C_n \subset F$ by 7.2.1 (2). Therefore, $C_n + C_n + G_0 \subset F$, and hence $\overset{\circ}{C}_n + \overset{\circ}{C}_n \subset \overset{\circ}{F}$. Consequently, $(\overset{\circ}{C}_n + \overset{\circ}{C}_n)^\# = \overset{\circ}{C}_n^\# + \overset{\circ}{C}_n^\# \subset \overset{\circ}{F}^\#$. Since C_n is symmetric; therefore, $\overset{\circ}{C}_n^\# = -\overset{\circ}{C}_n^\#$.

(d): This is checked by analogy.

(2): Let $F \in \text{In}(G_f)$. We will show that $F^\#$ is closed.

If $g^\# \notin F^\#$ then $(g + G_0) \cap F = \emptyset$, and so $(g + C_n) \cap F = \emptyset$ for some $n \in \mathbb{Z}$ (by ω^+ -saturation). Thus, $(g + C_{n-1} + G_0) \cap F = \emptyset$ by 7.2.1 (2), implying that $(g^\# + C_{n-1}^\#) \cap F^\# = \emptyset$. Hence, $(g^\# + \overset{\circ}{C}_{n-1}^\#) \cap F^\# = \emptyset$ because $\overset{\circ}{C}_{n-1}^\# \subset C_{n-1}$. This proves that $F^\#$ is closed since $\overset{\circ}{C}_{n-1}^\# \in \mathcal{U}$.

(3): It follows from 7.2.1 (2), that the canonical topology on $G^\#$ satisfies the first axiom of countability. It suffices so to show that every Cauchy sequence in $G^\#$ converges.

Let $(g_n^\#)_{n \in \mathbb{N}}$ be a Cauchy sequence in $G^\#$. By definition, to each $m \in \mathbb{Z}$ there is $\nu(m) \in \mathbb{N}$ such that $g_{n_1}^\# - g_{n_2}^\# \in \overset{\circ}{C}_m^\#$ for all $n_1, n_2 > \nu(m)$.

Consider the countable family

$$\Gamma := \{A_{m,n} : n > \nu(m), n, m \in \mathbb{N}\},$$

where

$$A_{m,n} = \{g : g_n - g \in C_{m+1}\},$$

and show that Γ has the finite intersection property. To this end, take

$$S := \{A_{m_1, n_1}, \dots, A_{m_k, n_k}\}$$

and choose $n > \max\{\nu(m_1), \dots, \nu(m_k)\}$. Then $g_{n_i}^\# - g_n^\# \in \overset{\circ}{C}_{m_i}^\#$. Therefore, $g_{n_i} - g_n + g \in \overset{\circ}{C}_{m_i}$ for some $g \in G_0$.

Thus, $g_{n_i} - g_n + g + G_0 \subset C_{m_i}$. Since $g + G_0 = G_0$, it follows that $g_{n_i} - g_n \in C_{m_i}$ ($i := 1, \dots, k$), and so $g_n \in \bigcap S$. By ω^+ -saturation, $\bigcap \Gamma$ is nonempty and there is some $g \in \bigcap \Gamma$. Considering that $C_{m+1} \subset \overset{\circ}{C}_{m+2}$ which is immediately from 7.2.1 (2), conclude that $g_n^\# \rightarrow g^\#$ as $n \rightarrow \infty$. \triangleright

Recall that an internal finite set A is *standardly finite* whenever the size $|A|$ of A is a standard natural; cf. 3.7.7. We also let $(*)$ stand for the condition holds

$$(\forall F_1, F_2 \in \text{In}_0(G_f))(F_1 \subset F_2 \rightarrow (\exists B \subset F_2)(|B| \in \mathbb{N} \wedge F_1 + B \supseteq F_2)).$$

7.2.3. Theorem. *The following hold:*

- (1) *The canonical topology on $G^\#$ is locally compact and separable if and only if $(*)$ is valid;*
- (2) *On assuming $(*)$, $G^\#$ is compact (discrete) if and only if G_f (respectively G_0) is an internal subgroup of G .*

\triangleleft (1): Assuming $(*)$, show that $G^\#$ is locally compact. It suffices to prove that $F^\#$ is compact for every $F \in \text{In}_0(G_f)$. The closure of $F^\#$ was shown in 7.2.2. Prove

that to each neighborhood U of zero in $G^\#$ there is a finite set $\{v_1, \dots, v_k\} \subset F^\#$ satisfying $\bigcup_{i=1}^k (v_i + U) \supset F^\#$.

By 7.2.1(2) we may choose $n \in \mathbb{Z}$ so that $C_n \subset F$ and $\overset{\circ}{C}_n^\# \subset U$. Then $C_{n-1} \subset \overset{\circ}{C}_n$, since $C_{n-1} + C_{n-1} \subset C_n$ and $G_0 \subset C_{n-1}$. The condition (*) implies that there is a finite set $B \subset F$ satisfying $C_{n-1} + B \supset F$. Then $\overset{\circ}{C}_n + B \supset F$ and $\overset{\circ}{C}_n^\# + B^\# \supset F^\#$. Of course, $B^\#$ is finite since B is standardly finite. The separability of $G^\#$ follows from its metrizable and the fact that $G^\# = \bigcup \{C_n^\# : n \in \mathbb{Z}\}$, with each $C_n^\#$ compact.

Suppose conversely that $G^\#$ is locally compact and separable. It is easy to see that there is some $n_0 \in \mathbb{Z}$ such that $C_n^\#$ is compact for all $n \leq n_0$. Show that $F^\#$ is compact for every $F \in \text{In}(G_f)$. Put $V := \overset{\circ}{C}_{n_0-1}^\#$. Then $\bigcup \{g^\# + V : g \in G_f\} = G^\#$. By separability, there is a sequence $\{g_n\}$ satisfying $G^\# = \bigcup_n (g_n^\# + V)$, and so

$$G_f = \bigcup_n (g_n + \overset{\circ}{C}_{n_0-1} + G_0) \subset \bigcup_n (g_n + C_{n_0}) \subset G_f.$$

Consequently, $F \subset \bigcup_n (g_n + C_{n_0})$ and, by ω_1 -saturation, there is a finite set $\{n_1, \dots, n_k\}$ satisfying $F \subset \bigcup_{i=1}^k g_{n_i} + C_{n_0}$. Thus, $F^\# \subset \bigcup_{i=1}^k g_{n_i} + C_{n_0}$, which also implies that $F^\#$ is compact. Suppose now that $G_0 \subset F_1 \subset F_2$ and let $n \leq n_0$ be such that $C_n \subset F_1$. There are g_1, \dots, g_k such that $g_1^\#, \dots, g_k^\# \in F_2^\#$ and $\bigcup_{i=1}^k g_i + \overset{\circ}{C}_{n-1} \supset F_2$. If $h_1, \dots, h_k \in F_2$ and $h_i - g_i \in G_0$ then $h_i + G_0 = g_i + G_0$. Hence,

$$F_2 + G_0 \subset \bigcup_{i=1}^k (g_i + \overset{\circ}{C}_{n-1} + G_0) \subset \bigcup_{i=1}^k (h_i + C_n) \subset \{h_1, \dots, h_k\} + F_1.$$

Putting $B := \{h_1, \dots, h_k\}$, we arrive at (*).

(2): Suppose now that $G^\#$ is compact. Given $F \in \text{In}_0(G_f)$, find $g_1, \dots, g_k \in G_f$ satisfying

$$G^\# = \bigcup_{i=1}^k (g_i^\# + \overset{\circ}{F}^\#) = \left(\bigcup_{i=1}^k (g_i + \overset{\circ}{F}) \right)^\# = \left(\bigcup_{i=1}^k (F + g_i) \right)^\# \quad (\overset{\circ}{F} \subset F \subset G_f).$$

Obviously, $K := \bigcup_{i=1}^k (F + g_i) \subset G_f$ is internal and $K^\# = G^\#$. Hence, $G_f = K + G_0 \subset K + F \subset G_f$. Thus, $G_f = K + F$ is an internal set too.

Suppose conversely that G_f is an internal subgroup of G . Show that $G^\#$ is compact on assuming (*), of course. It suffices to demonstrate that

$$(\forall n \in \mathbb{Z})(\exists B \subset G_f) \left(|B| \in \mathbb{N} \wedge G^\# = \bigcup_{g \in B} (g^\# + \overset{\circ}{C}_n^\#) \right).$$

By 7.2.1 (2), $C_{n-1} \subset \overset{\circ}{C}_n$ and by (*) there is a standardly finite B such that

$$G_f = B + C_{n-1} = B + \overset{\circ}{C}_n = \bigcup_{g \in B} g + \overset{\circ}{C}_n.$$

Now $G^\# = \bigcup_{g \in B} (g^\# + \overset{\circ}{C}_n^\#)$. We are done on omitting the simple proof of the last claim of (2). \triangleright

It is worth noting that, in the context of Theorem 7.2.3, $F^\#$ is compact for all $F \in \text{In}(G_f)$.

7.2.4. We now list a few auxiliary facts about the entities under study which we will use in the sequel.

(1) If $F \subset G_f$ then $g + \overset{\circ}{F} = (g + F)^\circ$, $j^{-1}(j(\overset{\circ}{F})) = \overset{\circ}{F}$, $(g + \overset{\circ}{F})^\# = g^\# + \overset{\circ}{F}^\#$, and $\overset{\circ}{F}^\# = {}^c(G_f - F)^\#$, where ${}^c(A)$ stands for the complement of a set A .

(2) If $F \in \text{In}(G_f)$ then $\overset{\circ}{F}^\#$ is open and $\{\overset{\circ}{F}^\# : F \in \text{In}(G_f)\}$ is a base for the canonical topology of $G^\#$.

It is worth noting that we will proceed in the context of Theorem 7.2.3 by implication.

(3) If $F_1, F_2 \in \text{In}_0(G_f)$ then $0 < \circ(|F_1|/|F_2|) < +\infty$.

\triangleleft By Theorem 7.2.3 there is a standardly finite B (i.e., $|B| \in \mathbb{N}$) such that $F_1 + B \supset F_1 \cup F_2 \supset F_2$. Then $|F_2| \leq |F_1 + B| \leq |F_1||B|$, and so $\circ(|F_2|/|F_1|) < +\infty$. Similarly, $\circ(|F_1|/|F_2|) < +\infty$, which completes the proof. \triangleright

(4) A hyperreal $\Delta \in {}^*\mathbb{R}_+$ is a *normalizing factor* for a triple (G, G_0, G_f) provided that $0 < \circ(\Delta|F|) < +\infty$ for all $F \in \text{In}_0(G_f)$.

It is immediately from (3) that if $F \in \text{In}_0(G_f)$ then $\Delta := |F|^{-1} \in {}^*\mathbb{R}_+$ is a normalizing factor. Consequently, there is a normalizing factor for each triple (G, G_0, G_f) satisfying the conditions of Theorem 7.2.3. It is also clear that, in case Δ is a normalizing factor, for Δ' to be a normalizing factor it is necessary and sufficient that $0 < \circ(\frac{\Delta}{\Delta'}) < +\infty$. Moreover, (3) shows that if Δ is a normalizing factor for (G, G_0, G_f) then $(G_f, S_{\Delta}^{G_f}, \nu_{\Delta}^{G_f})$ is a σ -finite subspace of the Loeb measure space $(G, S_{\Delta}, \nu_{\Delta})$.

For the sake of simplicity, we agree to write S instead of $S_{\Delta}^{G_f}$ and ν_{Δ} instead of $\nu_{\Delta}^{G_f}$.

(5) If $A \in S$ and $g \in G_f$ then $g + A \in S$ and $\nu_{\Delta}(g + A) = \nu_{\Delta}(A)$.

\triangleleft Obvious. \triangleright

(6) Let \mathcal{B} stand for the Borel σ -algebra of $G^\#$. Then $j^{-1}(B) \in S$ for all $B \in \mathcal{B}$.

◁ By (1) it suffices to show that $j^{-1}((G_f - F)^\#) \in S$ for $F \in \text{In}(G_f)$. We thus complete the proof by the chain of equalities

$$\begin{aligned} j^{-1}((G_f - F)^\#) &= G_f - F + G_0 = \bigcup_{n \in \mathbb{Z}} (C_n - F) + \bigcap_{m \in \mathbb{Z}} C_m, \\ \bigcup_{n \in \mathbb{Z}} (C_n - F) + \bigcap_{m \in \mathbb{Z}} C_m &= \bigcup_{n \in \mathbb{Z}} (C_n - F + \bigcap_{m \in \mathbb{Z}} C_m), \\ C_n - F + \bigcap_{m \in \mathbb{Z}} C_m &= \bigcap_{m \in \mathbb{Z}} (C_n - F + C_m). \end{aligned}$$

Observe that this final equality rests on ω^+ -saturation. ▷

We now define the measure μ_Δ on \mathcal{B} by putting

$$\mu_\Delta(B) := \nu_\Delta(j^{-1}(B)).$$

From (5) it is immediate that μ_Δ is an invariant measure, while μ_Δ is regular because $G^\#$ is separable. Thus, μ_Δ is a Haar measure on $G^\#$. We denote by L the completion of the σ -algebra \mathcal{B} with respect to μ_Δ . The extension of μ_Δ to L is again denoted by μ_Δ .

7.2.5. Theorem. *The following hold:*

- (1) *A subset A of $G^\#$ belongs to L if and only if $j^{-1}(A) \in S$.*
- (2) *$\mu_\Delta(B) = \nu_\Delta(j^{-1}(B))$ for all $B \in L$.*

◁ The fact that $j^{-1}(A) \in S$ ($A \in L$) and $\mu_\Delta(A) = \nu_\Delta(j^{-1}(A))$ hold follows immediately from the completeness of the Loeb measure ν_Δ .

It obviously suffices to prove the converse for $A \subset G^\#$ such that $j^{-1}(A) \subset F \in \text{In}_0(G_f)$. To show that $A \in L$ and $\mu_\Delta(A) = \nu_\Delta(j^{-1}(A))$ holds, it suffices to check that $\mu_{\text{in}}(A) = \mu_{\text{out}}(A) = \nu_\Delta(j^{-1}(A))$, where $\mu_{\text{in}}(A)$ and $\mu_{\text{out}}(A)$ are the inner and outer Haar measures of A .

Since $\nu_\Delta(F)$ is limited and, hence, so is $\nu_\Delta(j^{-1}(A))$; to each standard $\varepsilon > 0$ there is an internal set $\mathcal{D} \subset j^{-1}(A)$ satisfying $\nu_\Delta(\mathcal{D}) \geq \nu_\Delta(j^{-1}(A)) - \varepsilon$. Since $\mathcal{D}^\# \subset A$ and $\mathcal{D}^\#$ is closed; therefore, $\mu_{\text{in}}(A) \geq \nu_\Delta(j^{-1}(A))$.

Suppose now that $H \in \text{In}_0(G_f)$ is such that $A \subset \overset{\circ}{H}^\#$ (for example, we may take $F + F$ as H) and put $B := \overset{\circ}{H}^\# - A$. Then $j^{-1}(B) = \overset{\circ}{H} - j^{-1}(A)$ and $\nu_\Delta(j^{-1}(A)) + \nu_\Delta(j^{-1}(B)) = \nu_\Delta(\overset{\circ}{H}) = \mu_\Delta(\overset{\circ}{H}^\#)$, since $j^{-1}(\overset{\circ}{H}^\#) = \overset{\circ}{H}$ by 7.2.4(1). This implies that $\mu_{\text{in}}(A) + \mu_{\text{in}}(B) \geq \mu_\Delta(\overset{\circ}{H}^\#)$. On the other hand, it follows from

the regularity of μ_Δ that

$$\begin{aligned} \mu_\Delta(\overset{\circ}{H^\#}) &= \sup\{\mu_\Delta(E) : E \subset \overset{\circ}{H^\#}, E \text{ is closed}\} \\ &\geq \sup\{\mu_\Delta(C) + \mu_\Delta(\mathcal{D}) : C \subset A, \mathcal{D} \subset B, C \text{ and } \mathcal{D} \text{ are closed}\} \\ &= \sup\{\mu_\Delta(C) : C \subset A, C \text{ is closed}\} \\ &\quad + \sup\{\mu_\Delta(\mathcal{D}) : \mathcal{D} \subset B, \mathcal{D} \text{ is closed}\} \\ &= \mu_{\text{in}}(A) + \mu_{\text{in}}(B). \end{aligned}$$

Thus, $\mu_{\text{in}}(A) + \mu_{\text{in}}(B) = \mu_\Delta(\overset{\circ}{H^\#})$. Now, if $\varepsilon > 0$ then there is a closed subset C of B satisfying $\mu_\Delta(C) \geq \mu_{\text{in}}(B) - \varepsilon$. Since A lies in the open set $\overset{\circ}{H^\#} - C$; therefore,

$$\begin{aligned} \mu_{\text{out}}(A) &\leq \mu_\Delta(\overset{\circ}{H^\#} - C) = \mu_\Delta(\overset{\circ}{H^\#}) - \mu_\Delta(C) \\ &\leq \mu_\Delta(\overset{\circ}{H^\#}) - \mu_{\text{in}}(B) + \varepsilon = \mu_{\text{in}}(A) + \varepsilon. \end{aligned}$$

Thus, $\mu_{\text{out}}(A) = \mu_{\text{in}}(A)$, implying easily that $\mu_\Delta(A) = \nu_\Delta(j^{-1}(A))$. \triangleright

7.2.6. Theorem 7.2.5 shows that $j : G_f \rightarrow G^\#$ is a measure-preserving mapping. If $f : G^\# \rightarrow \mathbb{R}$ is a μ_Δ -measurable function then $f \circ j : G_f \rightarrow \mathbb{R}$ is a ν_Δ -measurable function. A lifting φ of the latter is called a *lifting* of f . Thus, an internal function $\varphi : G \rightarrow {}^*\mathbb{R}$ is a lifting of f whenever $f(g^\#) = \overset{\circ}{\varphi}(g)$ for ν_Δ -almost all $g \in G_f$.

If $\varphi \in \mathcal{S}_p(G_f)$ then φ is an \mathcal{S}_p -integrable lifting of f . A somewhat more precise expression “ φ is an $\mathcal{S}_{p,\Delta}$ -lifting of f ” is also in common parlance.

A function f belongs to $L_p(\mu_\Delta)$, with $p \in [1, \infty]$, if and only if f has an $\mathcal{S}_{p,\Delta}$ -integrable lifting $\varphi : G \rightarrow {}^*\mathbb{R}$.

Furthermore,

$$\int_{G^\#} f d\mu_\Delta = \overset{\circ}{\left(\Delta \sum_{g \in G} \varphi(g) \right)}$$

and

$$\int_{G^\#} |f|^p d\mu_\Delta = \overset{\circ}{\left(\Delta \sum_{g \in G} |\varphi(g)|^p \right)}$$

for all $p \in [1, \infty]$

7.2.7. Let $G^\wedge := \widehat{G}$ be the internal character group of G . Since G is hyperfinite; by transfer we infer that G^\wedge is isomorphic to G and so G^\wedge is an internal hyperfinite abelian group.

Following [402], we represent the group S^1 (more exactly, the underlying set of S^1 which is the unit circle) as the interval $[-\frac{1}{2}, \frac{1}{2}]$ with addition modulo 1. We take in S^1 the countable family $\{\Lambda_k : k := 1, 2, \dots\}$ of neighborhoods about zero, with $\Lambda_k := (-\frac{1}{3k}, \frac{1}{3k})$. In the sequel we need a few auxiliary facts.

(1) If $\gamma \in \Lambda_1, 2\gamma \in \Lambda_1, \dots, k\gamma \in \Lambda_1$ then $\gamma \in \Lambda_k$.

◁ Obvious; cf. [402]. ▷

We now define the two external subgroups $H_0 \subset H_f$ of G^\wedge by the formulas

$$\begin{aligned}\alpha \in H_0 &\leftrightarrow (\forall g \in G_f)(\alpha(g) \approx 0), \\ \alpha \in H_f &\leftrightarrow (\forall g \in G_0)(\alpha(g) \approx 0).\end{aligned}$$

We will also use the countable family $\{W(C_n, \Lambda_k) : n \in \mathbb{Z}, k \in \mathbb{N}\}$ of internal subsets of $W(C_n, \Lambda_k) \subset G^\wedge$ such that

$$\alpha \in W(C_n, \Lambda_k) \leftrightarrow (\forall g \in C_n)(\alpha(g) \in \Lambda_k).$$

Given $F \in \text{In}(G_f)$, we define $W(F, \Lambda_k)$ similarly.

(2) $H_0 = \bigcap_{n,k} W(C_n, \Lambda_k)$ and $H_f = \bigcup_n W(C_n, \Lambda_1)$.

◁ The first equality is rather obvious and so we prove the second.

Suppose that $\alpha \in W(C_n, \Lambda_1)$, and $m \in \mathbb{Z}$ is such that $kC_m \subset C_n$. In this event if $g \in C_m$ then $\alpha(g), 2\alpha(g), \dots, k\alpha(g) \in \Lambda_1$, and so $\alpha(g) \in \Lambda_k$ by (1). Consequently, $\alpha(g) \in \Lambda_k$ for all k and $g \in G_0$, implying that $\alpha(g) \approx 0$ and $\alpha \in H_f$.

Conversely, let $\alpha \in H_f$, and assume that $\alpha \notin W(C_n, \Lambda_1)$ for all n . Then to each n there is g in C_n satisfying $(|\alpha(g)| \geq \frac{1}{3})$. By ω_1 -saturation, there is some g in G_0 such that $(|\alpha(g)| \geq \frac{1}{3})$, which contradicts the membership $\alpha \in H_f$. ▷

Thus, the triple (G^\wedge, H_0, H_f) satisfies the same conditions as (G, G_0, G_f) , and so we may define the canonical topology on $G^{\wedge\#} = H_f/H_0$. It follows from Theorem 7.2.3 that $G^{\wedge\#}$ is complete (with respect to the corresponding uniformity). If $\alpha \in H_f$ then we let $\alpha^\#$ stand for the image of α in $G^{\wedge\#}$.

(3) Let the internal function $f : G \rightarrow {}^*\mathbb{C}$ be such that $f(g_1) \approx f(g_2)$ whenever g_1, g_2 and $g_1 - g_2 \in G_0$ while ${}^\circ|f(g)| < +\infty$ for $g \in G_f$. Then the function $\tilde{f} : G^\# \rightarrow \mathbb{C}$, with $\tilde{f}(g^\#) := {}^\circ f(g)$ for all $g \in G_f$, is uniformly continuous on $G^\#$.

◁ Obvious. ▷

We now define the mapping $\psi : G^{\wedge\#} \rightarrow G^{\#\wedge}$ by putting

$$\psi(\alpha^\#)(g^\#) := {}^\circ(\alpha(g))$$

for all $\alpha \in H_f$ and $g \in G_f$. The membership $\psi(\alpha^\#) \in G^{\#\wedge}$ follows from (3). Also, it is an easy matter to show that ψ is soundly defined monomorphism.

7.2.8. Theorem. *The mapping $\psi : G^{\wedge\#} \rightarrow \psi(G^{\wedge\#}) \subset G^{\#\wedge}$ is a topological isomorphism.*

◁ Recall (see, for instance, [402]), that the topology on the dual group $G^{\#\wedge}$ of $G^\#$ is determined by the neighborhood base of zero comprising all sets of the shape $\mathcal{W}(\mathcal{F}, \Lambda_k)$, where

$$\mathcal{W}(\mathcal{F}, \Lambda_k) := \{h \in G^{\#\wedge} : (\forall \xi \in \mathcal{F})(h(\xi) \in \Lambda_k)\},$$

with \mathcal{F} a compact set in $G^\#$ and $k \in \mathbb{N}$.

It is now easy to see that $\{\mathcal{W}(C_n^\#, \Lambda_k) : n \in \mathbb{Z}, k \in \mathbb{N}\}$ is also a neighborhood base of zero in $G^{\#\wedge}$.

The easy inclusions

$$\begin{aligned} \circ\psi(\overset{\circ}{W}(C_n, \Lambda_{k+1})) &\subset \mathcal{W}(C_n^\#, \Lambda_k), \\ \psi^{-1}(\mathcal{W}(C_n^\#, \Lambda_{k+1})) &\subset W(C_n, \Lambda_k)^\# \end{aligned}$$

imply that ψ and ψ^{-1} are continuous. ▷

7.2.9. Two corollaries of Theorem 7.2.8 are now in order.

(1) *The image $\psi(G^{\wedge\#})$ of ψ is a closed subgroup of $G^{\#\wedge}$.*

◁ This is immediate since $G^{\wedge\#}$ is complete and ψ is uniformly continuous. ▷

(2) *$G^{\#\wedge}$ is a separable locally compact group.*

Conjecture. *If (G, G_0, G_f) satisfies the conditions of Theorem 7.2.3 then $\psi : G^{\wedge\#} \rightarrow G^{\#\wedge}$ is a topological isomorphism, i.e.,*

$$\psi(G^{\wedge\#}) = G^{\#\wedge}.$$

7.2.10. For brevity we let $H := G^\wedge$ and identify $H^\#$ with $\psi(H^\#)$ by putting $h^\#(g^\#) := \circ(h(g))$ for all $h \in H_f$ and $g \in G_f$. Thus, $G^{\wedge\#} = H^\#$ and $H^\#$ is a closed subgroup of $G^{\#\wedge}$. Assign

$$\begin{aligned} G'_0 &:= \{g \in G : (\forall \alpha \in H_f)(\alpha(g) \approx 0)\}, \\ G'_f &:= \{g \in G : (\forall \alpha \in H_0)(\alpha(g) \approx 0)\}. \end{aligned}$$

Obviously, $G'_0 \supset G_0$ and $G'_f \supset G_f$. Put $G^{\#\prime} := G'_f/G'_0$.

(1) **Theorem.** *In the above context,*

$$H^\# = G^{\#\wedge} \leftrightarrow G'_0 \cap G_f = G_0.$$

◁ Since $G = H^\wedge$, we may apply Theorem 7.2.8 and 7.2.9(1) to the couple (G'_0, G'_f) and infer that $G^{\#'}$ is a closed subgroup of $H^{\#\wedge}$. Since $H^\#$ is a closed subgroup of $G^{\#\wedge}$, the Pontryagin Duality Theorem (see [402]) yields

$$H^{\#\wedge} = G^\# / \text{Ann}(H^\#),$$

where $\text{Ann}(H^\#) := \{\xi \in G^\# : (\forall h \in H^\#)(h(\xi) = 0)\}$.

Let $g^{\#'}$ stand for the canonical image of $g \in G'_f$ in $G^{\#'}$. Since $g^{\#'}$ is a character of $H^\#$, there is an element $g_1 \in G_f$ satisfying $g^{\#'}(h^\#) = g_1^\#(h^\#)$ for all $h \in H_f$. Hence, $h(g - g_1) \approx 0$ for all $h \in H_f$, implying that $g - g_1 \in G'_0$. Therefore, $(\forall g \in G'_f)(\exists g_1 \in G_f)(g - g_1 \in G'_0)$, and so $G^{\#'} = G_f/G_f \cap G'_0$.

Suppose now that $G_f \cap G'_0 = G_0$. Then $G^{\#'} = G^\#$, and so $\text{Ann}(H^\#) = 0$ since $H^{\#\wedge} = G^\# / \text{Ann}(H^\#)$. Hence, $H^\# = G^{\#\wedge}$.

Suppose conversely that $H^\# = G^{\#\wedge}$, i.e., $H^{\#\wedge} = G^{\#\wedge\wedge}$ and so $H^{\#\wedge} = G^\#$. Then $\text{Ann}(H^\#) = 0$. If, nevertheless, $g \in G'_0 \cap G_f - G_0$ then $g^\# \in \text{Ann}(H^\#)$ and $g^\# \neq 0$, which would be a contradiction. ▷

$$(2) \quad G^{\#'} = H^{\#\wedge}.$$

◁ Since $G_f \subset G'_f$, the group $H'_0 = \{h \in H : (\forall g \in G'_f)(h(g) \approx 0)\}$ lies in H_0 . We are done on applying (1). ▷

7.2.11. Let S^1 stand as usual for the unit circle. If $\chi : G \rightarrow {}^*S^1$ is an internal character of G satisfying $\chi|_{G_0} \approx 0$, then there is some character $\tilde{\chi} : G^\# \rightarrow S^1$ such that $\tilde{\chi}(g^\#) = {}^\circ\chi(g)$ for all $g \in G_f$. We note that the equality $G^{\#\wedge} = G^{\wedge\#}$ means that every character $h : G^\# \rightarrow S^1$ has the shape of $\tilde{\chi}$. We will derive some sufficient conditions for this equality to hold. We start with an auxiliary proposition.

(1) *If K is an internal hyperfinite abelian group, and $X : K \rightarrow S^1$ is an internal character of K satisfying $X(g) \approx 1$ for all $g \in K$ then $X \equiv 1$.*

◁ Let $|K| = N$. Then $X(g)^N = 1$, i.e., $X(g) = \exp(2\pi i m_X(g)/N)$. Obviously, $m_X : G \rightarrow \mathbb{Z}/N\mathbb{Z}$ is a group homomorphism. Therefore, $m_X(G)$ is a cyclic subgroup of $\mathbb{Z}/N\mathbb{Z}$, and so there is d divisible by N such that $Nm_X(G) = \{kd : 0 \leq k < N/d\}$.

If N/d is even then, putting $k := N/2d$, we find that

$$\exp(2\pi i kd/N) = \exp(\pi i) = -1,$$

which contradicts the hypothesis. If N/d is odd then we put $k := (N/d - 1)/2$. In this event $\exp(2\pi i kd/N) = -\exp(-\pi id/N) \approx -1$ whenever $d/N \approx 0$. If $d/N \not\approx 0$ then N/d is some standard number, say, m and $\exp(2\pi i kd/N) = -\exp(-\pi i/m) \not\approx 1$ for $m \neq 1$. Thus, $N = d$, and we may conclude that $m_X(G) = 0$, i.e., $X \equiv 1$. ▷

$$(2) \quad \text{If } G^\# \text{ is a discrete or compact group then } G^{\#\wedge} = G^{\wedge\#}.$$

◁ Let $G^\#$ be a discrete group. Then by the second part of Theorem 7.2.3 G_0 is an internal subgroup of G . By (1), $H_f = \{h \in G^\wedge : (\forall g \in G_0)(h(g) = 0)\}$. Hence, H_f is an internal subgroup of $H := G^\wedge$. Moreover, $G'_0 := \{g \in G : (\forall h \in H_f)(h(g) = 0)\}$ is an internal subgroup of G . Further, $H_f = \text{Ann}(G_0)$ and $G'_0 = \text{Ann}(H_f)$. By transfer and the theorem on duality of annihilators, $G_0 = G'_0$. The case of a compact group $G^\#$ (that is, of an internal group G_f) is settled in the next section. ▷

(3) *If there is a subgroup $K \in \text{In}_0(G_f)$ then $G^{\#\wedge} = G^{\wedge\#}$.*

◁ Note first that the triples (G_0, K, G) and (K, G_f, G) satisfy all conditions of Theorem 7.2.3. Hence, $K^\# = K/G_0$ is a compact subgroup of $G^\#$, and G_f/K is a discrete group. Therefore, we may apply (2) to these groups. We show now that $G'_0 \cap G_f = G_0$. If not, then $g_0 \in G'_0 \cap G_f - G_0$, and two cases are possible.

(a): $g_0 \in K$: By (2) there is an internal character $X : K \rightarrow {}^*S^1$ such that $X|_{G_0} \approx 0$ and $X(g_0) \not\approx 0$. By transfer X may be extended to some internal character $\varkappa \in H_f$. Further, $\varkappa(g_0) \not\approx 0$, contradicting the fact that $g_0 \in G'_0$.

(b): $g_0 \notin K$: Again by (2) there is an internal character $h : G \rightarrow {}^*S^1$ such that $h|_K \approx 1$ and $h(g_0) \not\approx 1$. Thus, $h \in H_f$ contradicting the membership $g_0 \in G'_0$.

Note that both cases use the fact that the characters of a locally compact abelian group separate its points. ▷

In the sequel we let (G, G_0, G_f) stand for a triple of groups of the environment of Theorem 7.2.3.

7.2.12. Theorem. *The group $G^\#$ has an open compact subgroup if and only if there is an internal subgroup $K \in \text{In}_0(G_f)$. Furthermore, $K^\#$ is an open compact subgroup of $G^\#$.*

◁ If $K \in \text{In}_0(G_f)$, then $K + G_0 = K$, which implies that $K^\#$ is an open set. The compactness of $K^\#$ is established in 7.2.3. Obviously, if K is a subgroup of G then $K^\#$ is a subgroup of $G^\#$.

Conversely, suppose that $U \subset G^\#$ is an open compact subgroup. We show that $j^{-1}(U)$ is an internal set. Let $F \in \text{In}_0(G_f)$ be such that $F^\# \subset U$. For example, we may take C_{n-1} to be F if $\overset{\circ}{C}_n^\# \subset U$ (see 7.2.1(2)). Such a C_n exists because U is open.

Since U is compact, there is $g_1^\#, \dots, g_n^\# \in U$ satisfying $U = \bigcup_{i=1}^n (g_i^\# + \overset{\circ}{F}^\#)$. Since U is a subgroup and $F^\# \subset U$; therefore, $U = \bigcup_{i=1}^n (g_i^\# + F^\#)$. Hence, $j^{-1}(U) = (\bigcup_{i=1}^n g_i + \overset{\circ}{F}) \subset \bigcup_{i=1}^n g_i + F \subset j^{-1}(U)$, which completes the proof. ▷

7.2.13. Let Δ be a normalizing factor of (G, G_0, G_f) (cf. 7.2.4(4)). As above in 7.1.1, put $\widehat{\Delta} := (\Delta|G|)^{-1}$. Recall that $\mathcal{L}_{2,\Delta}(G)$ is an internal hyperfinite-

dimensional space in the space ${}^*\mathbb{C}^G$ with the inner product

$$(\varphi, \psi)_\Delta := \Delta \sum_{g \in G} \varphi(g) \overline{\psi(g)}$$

for all internal $\varphi, \psi \in {}^*\mathbb{C}^G$. The $\mathcal{L}_{2, \widehat{\Delta}}(\widehat{G})$ space is defined similarly.

The *discrete Fourier transform* $\Phi_\Delta^G : \mathcal{L}_{2, \Delta}(G) \rightarrow \mathcal{L}_{2, \widehat{\Delta}}(\widehat{G})$ is defined as

$$\Phi_\Delta^G(\varphi)(\chi) := (\varphi, \chi)_\Delta \quad (\varphi \in \mathcal{L}_{2, \Delta}(G), \chi \in \widehat{G}).$$

Clearly, the discrete Fourier transform Φ_Δ^G preserves the inner product.

A triple of groups (G, G_0, G_f) is *admissible* provided that the following hold:

- (1) $G^{\#\wedge} = G^{\wedge\#}$;
- (2) $\widehat{\Delta}$ is a normalizing factor for the triple (\widehat{G}, H_0, H_f) in 7.2.7;
- (3) Φ_Δ^G is a hyperapproximant to the Fourier transform $\mathcal{F}_\Delta^{G^\#} : L_2(\mu_\Delta) \rightarrow L_2(\mu_{\widehat{\Delta}})$, defined as

$$\mathcal{F}_\Delta^{G^\#}(f)(\varkappa) := \int f(g) \overline{\varkappa(g)} d\mu(g) \quad (\varkappa \in G^{\#\wedge}).$$

Theorem 7.1.2 shows that the triple (G, G_0, G_f) of 7.1.1 is admissible.

7.2.14. Theorem. *If there is a subgroup K belonging to $\text{In}_0(G_f)$ then the group (G, G_0, G_f) is admissible.*

◁ Show first that $\widehat{\Delta}$ is a normalizing factor for \widehat{G} . To this end, consider the subgroup $K^\perp := \{\chi \in \widehat{G} : \chi|_K = 1\}$. It follows from 7.2.11 (1) that $H_0 \subset K^\perp \subset H_f$. We are thus left with demonstrating that $0 < \circ(\widehat{\Delta}|K^\perp|) < +\infty$. To this end, note that $\widehat{K} = \widehat{G}/K^\perp$ (see 7.2.4 (3)). Hence, $|K^\perp| = |\widehat{G}|/|\widehat{K}| = |G|/|K|$ and so $\widehat{\Delta}|K^\perp| = (\Delta|K|)^{-1}$, whereas $0 < \circ(\Delta|K|) < +\infty$ because Δ is a normalizing factor for (G, G_0, G_f) and $K \in \text{In}_0(G_f)$.

We know that $K^\#$ is an open compact subgroup of $G^\#$ by Theorem 7.2.12. Therefore, it is discrete; and $G^\#/K^\#$ is countable since $G^\#$ is separable. Assume that $\{\xi_k : k \in \mathbb{N}\}$ is complete system of representatives of the cosets of $G^\#/K^\#$ and $\varkappa \in K^{\#\wedge}$. Define the function $f_{k\varkappa} : G^\# \rightarrow \mathbb{C}$ as follows: Given $\eta \in G^\#$, put

$$f_{k\varkappa}(\eta) := \begin{cases} 0, & \text{if } \eta \notin \xi_k + K^\#, \\ \varkappa(\eta - \xi_k), & \text{if } \eta \in \xi_k + K^\#. \end{cases}$$

Let $\mathfrak{M} := \{f_{k\varkappa} : k \in \mathbb{N}, \varkappa \in K^{\#\wedge}\}$. Since linear combinations of characters are dense in $L_2(K^\#)$ and $L_2(G^\#) = \bigoplus_{k=0}^{\infty} L_2(\xi_k + K^\#)$, the linear span of \mathfrak{M} is dense in $L_2(G^\#)$, and we may apply 6.5.1 (2).

Let $\{x_k : k \in \mathbb{N}\} \subset G_f$ be such that $x_k^\# = \xi_k$ for all $k \in \mathbb{N}$. Choose $\varkappa_0 \in K^{\#\wedge}$ and some $\chi_0 \in \widehat{K}$ so that $\chi_0|_{G_0} \approx 1$ and $\tilde{\chi}_0 = \varkappa_0$ (cf. Theorem 7.2.11 (2)). Define the internal function $\varphi_{k\chi_0} : G \rightarrow {}^*\mathbb{C}$ by putting

$$\varphi_{k\chi_0}(y) := \begin{cases} 0, & \text{if } y \notin x_k + K, \\ \chi_0(y - x_k), & \text{if } y \in x_k + K, \end{cases}$$

for all $y \in G$.

Since $K + G_0 = K$, it follows easily that ${}^\circ\varphi_{k\chi_0}(y) = f_{k\varkappa_0}(y^\#)$ for all $y \in G_f$; in other words, $\varphi_{k\chi_0}$ is a lifting of the function $f_{k\varkappa_0}$. Since $\varphi_{k\chi_0}$ is bounded and supported in the internal set $x_k + K \subset G_f$; therefore, $\varphi_{k\chi_0} \in \mathcal{S}_{2,\Delta}(G)$. By 6.5.1 (2) it suffices to show that $\Phi_\Delta^G(\varphi_{k\chi_0})$ is an $\mathcal{S}_{2,\widehat{\Delta}}$ -integrable lifting $\mathcal{F}_\Delta^{G^\#}(f_{k\varkappa_0})$. Given $\varkappa \in G^{\#\wedge}$ and $\chi \in G^\wedge$, straightforward calculation yields

$$\begin{aligned} \mathcal{F}_\Delta^{G^\#}(f_{k\varkappa_0})(\varkappa) &= \begin{cases} \overline{\varkappa(\xi_k)}\mu_\Delta(K^\#), & \text{if } \varkappa|_{K^\#} = \varkappa_0, \\ 0, & \text{if } \varkappa|_{K^\#} \neq \varkappa_0; \end{cases} \\ \Phi_\Delta^G(\varphi_{k\chi_0})(\chi) &= \begin{cases} \overline{\chi(x_k)}\Delta|K|, & \text{if } \chi|_K = \chi_0, \\ 0, & \text{if } \chi|_K \neq \chi_0. \end{cases} \end{aligned}$$

Since $j^{-1}(K^\#) = K$; therefore, $\mu_\Delta(K^\#) = {}^\circ(\Delta|K|)$. If $\varkappa := \tilde{\chi}$ then it obviously follows that $\varkappa|_{K^\#} = \varkappa_0 \leftrightarrow \chi|_K = \chi_0$ (cf. 7.2.11 (1)).

It is now clear that ${}^\circ\Phi_\Delta^G(\varphi_{k\chi_0}) = \mathcal{F}_\Delta^{G^\#}(f_{k\varkappa_0})(\tilde{\varkappa})$ and $\Phi_\Delta^G(\varphi_{k\chi})$ is a lifting of $\mathcal{F}_\Delta^{G^\#}(f_{k\varkappa})$, because $\tilde{\chi}$ coincides with $\chi^\#$.

The internal function $\Phi_\Delta^G(\varphi_{k\chi_0})$ is bounded and supported in the internal set $\{\chi \in G^\wedge : \chi|_K = \chi_0\} \subset H_f$, and so $\Phi_\Delta^G(\varphi_{k\chi_0}) \in \mathcal{S}_{2,\widehat{\Delta}}(H_f)$. \triangleright

7.3. The Case of a Compact Nonstandard Hull

This section deals with a group G such that $G^\#$ is a compact group.

7.3.1. Assume that G is an internal hyperfinite group, and G_0 is a subgroup of G presented as the intersection of a countable family of internal sets and such that to each F satisfying $G_0 \subset F \subset G$, there is a standardly finite subset B of G enjoying the property $F + B = G$. In this event $G^\#$ is a compact group according to Theorem 7.2.3.

An internal function $\varphi : G \rightarrow {}^*\mathbb{C}$ is *S-continuous* provided that $\varphi(g_1) \approx \varphi(g_2)$ for all $g_1, g_2 \in G$ satisfying $g_1 - g_2 \in G_0$.

By 7.2.7 (3), if $\varphi : G \rightarrow {}^*\mathbb{C}$ is a pointwise limited *S-continuous* function, the former attribute meaning that ${}^\circ|\varphi(g)| < +\infty$ for all $g \in G$; then the function $\tilde{\varphi} : G^\# \rightarrow \mathbb{C}$, acting by the rule $\tilde{\varphi}(g^\#) = {}^\circ\varphi(g)$ for all $g \in G$, is continuous.

We will show in 7.3.4 that each continuous function from $G^\#$ to \mathbb{C} admits the above representation. To this end, two auxiliary facts are in order.

Denote by $CS(G)$ the set of all pointwise limited S -continuous internal functions $\varphi : G \rightarrow {}^*\mathbb{C}$. Clearly, $CS(G)$ is an external subalgebra of the internal algebra ${}^*\mathbb{C}^G$. Put $\mathfrak{G} := \{\tilde{\varphi} : \varphi \in CS(G)\}$. Note that \mathfrak{G} is a subalgebra of $C(G^\#)$. The following result is a “discrete” version of the Urysohn Theorem.

7.3.2. *If $x, y \in G$ and $x - y \notin G_0$ then there is some φ in $CS(G)$ such that $\varphi(x) = 0$ and $\varphi(y) = 1$.*

\triangleleft By ω^+ -saturation, $y \notin x + C_k$ for some $k \in \mathbb{Z}$ (cf. 7.2.1 (2)). Given $n \in \mathbb{Z}$, put $V_0 := C_k$ and $V_n := C_{k-n}$. Then $(V_n)_{n \in \mathbb{N}}$ is a sequence of symmetric internal sets satisfying $V_{n+1} + V_{n+1} \subset V_n$. By extension, there is an internal sequence $(B_n)_{n \in {}^*\mathbb{N}}$ of symmetric subsets of G such that $B_{n+1} + B_{n+1} \subset B_n$ for all $n \in {}^*\mathbb{N}$ and $B_k = V_k$ for all standard $k \in \mathbb{N}$.

Distinguish an infinite $N \in {}^*\mathbb{N}$. Suppose that $0 \leq m < n \leq N$, and let $V_{m,n} := B_{m+1} + \cdots + B_n$. By induction on n it is immediately from the inclusion $B_{n+1} + B_{n+1} \subset B_n$ that $V_{m,n} + B_n \subset B_m$. Consider the rationals of the shape

$$r = \frac{a_1}{2} + \frac{a_2}{2^2} + \cdots + \frac{a_n}{2^n}, \quad a_i \in \{0, 1\}.$$

Given $a \in \{0, 1\}$, put

$$B^a := \begin{cases} B, & \text{if } a = 1, \\ 0, & \text{if } a = 0. \end{cases}$$

Assign $W_r := B_1^{a_1} + B_2^{a_2} + \cdots + B_n^{a_n}$. Then $W_r \subset B_1 + B_2 + \cdots + B_n = V_{0N} \subset B_0 = V_0 = C_k$. Hence, $y \notin x + B_0$ and so $y \notin x + W_r$. Clearly, $W_r \subset W_{r'}$ for $r < r'$. Define the internal function $\varphi : G \rightarrow {}^*\mathbb{R}$ by the rule $\varphi(g) := \min\{r : g \in x + W_r\}$. If $g \notin W_r + x$ for all rationals r of the above shape then $\varphi(g) = 1$ and, in particular, $\varphi(y) = 1$. Since $x \in x + 0$, it follows that $\varphi(x) = 0$.

Show now that $|\varphi(u) - \varphi(v)| \leq \frac{1}{2^{k-1}}$ for $u - v \in B_k$ and $k \in [0, N]$. This will imply that φ is S -continuous for if $u - v \in G_0$ then $u - v \in B_k$ for every standard k . Since B_k is symmetric, we may assume that $\varphi(u) < \varphi(v)$ and observe that $\varphi(v) - \varphi(u) \leq \frac{1}{2^{k-1}}$. Note also that $\varphi(u) < 1$ because $\max \varphi = 1$. Assume that $q \in \{1, 2, \dots, 2^k\}$, with $k \leq N$, satisfies $\frac{q-1}{2^k} \leq \varphi(u) < \frac{q}{2^k}$. If $q = 2^k$ or $q = 2^k - 1$ then $1 - \varphi(u) \leq \frac{1}{2^{k-1}}$ and $\varphi(v) - \varphi(u) \leq \frac{1}{2^{k-1}}$ because $\varphi(v) \leq 1$.

Suppose now that $q < 2^{k-1} - 1$ and $r = \frac{q}{2^k}$. Then $\varphi(u) < r$ which amounts to $u \in W_r + x$ by the definition of φ . Since $v - u \in B_k$; therefore, $v \in W_r + B_k + x$. In this event $r = \frac{a_1}{2} + \cdots + \frac{a_k}{2^k}$, with $a_m \in \{0, 1\}$, and, moreover, there is some m satisfying $a_m = 0$ because $q < 2^k - 1$. Choose the greatest m with this property. Considering the inclusion $B_s + B_s \subset B_{s-1}$, infer

$$W_r + B_k = B_1^{a_1} + \cdots + B_{i+1}^{a_{i+1}} + \cdots + B_k^{a_k} + B_k \subset B_1^{a_1} + \cdots + B_{i-1}^{a_{i-1}} + B_i.$$

At the same time, $r' := r + \frac{1}{2^k} = \frac{a_1}{2} + \dots + \frac{a_{i-1}}{2^{i-1}} + \frac{1}{2^k}$. Thus, $W_r + B_k = W_{r'}$, and so $v \in W_{r'} + x$ and $\varphi(v) \leq r + \frac{1}{2^k}$. Finally, $r - \frac{1}{2^k} \leq \varphi(u) < \varphi(v) \leq r + \frac{1}{2^k}$, implying that $\varphi(v) - \varphi(u) \leq \frac{1}{2^{k-1}}$. \triangleright

7.3.3. *The subalgebra \mathfrak{G} is uniformly closed in $C(G^\#)$.*

\triangleleft Note first that, given $\psi \in CS(G)$, we obviously have

$$\sup\{\tilde{\psi}(\xi) : \xi \in G^\#\} = \circ \max\{\psi(g) : g \in G\}.$$

Thus, if $\{\tilde{\varphi}_n\}$ is a sequence of functions converging in $C(G^\#)$ to some function f then $\circ \max\{|\varphi_n(g) - \varphi_m(g)| : g \in G\} \rightarrow 0$ as $n, m \rightarrow \infty$. Therefore, there is a standard function N satisfying

$$\max\{|\varphi_{n_1}(g) - \varphi_{n_2}(g)| : g \in G\} < \frac{1}{m}$$

for all $n_1, n_2 > N(m)$. We now consider the family of internal sets

$$\{\{\varphi : \max\{|\varphi_n(g) - \varphi(g)| : g \in G\} < 1/m\} : n > N(m)\}.$$

which possesses the finite intersection property by the choice of N .

By ω^+ -saturation, there is an internal function $\varphi : G \rightarrow {}^*\mathbb{C}$ satisfying

$$\circ \max\{|\varphi_n(g) - \varphi(g)| : g \in G\} \xrightarrow{n \rightarrow \infty} 0.$$

It is clear that φ is pointwise limited and S -continuous. Hence, $\tilde{\varphi}_n \rightarrow \tilde{\varphi}$ as $n \rightarrow \infty$. Finally, $\tilde{\varphi} = f$. \triangleright

7.3.4. Theorem. *Each continuous function $f : G^\# \rightarrow \mathbb{C}$ has the shape $f = \tilde{\varphi}$ with $\varphi : G \rightarrow {}^*\mathbb{C}$ some S -continuous and pointwise limited function.*

\triangleleft To prove, it suffices to show that \mathfrak{G} is a uniformly closed subalgebra separating the points of $G^\#$. The closure is contained in 7.3.3, while the separatedness is immediate from 7.3.2.

Indeed, assume that $\xi, \eta \in G^\#$ and $\xi \neq \eta$. If $\xi = x^\#$ and $\eta = y^\#$ then $x - y \notin G_0$ and by 7.3.2 there is some φ in $CS(G)$ satisfying $\varphi(x) = 0$ and $\varphi(y) = 1$. But then $\tilde{\varphi}(\xi) = 0$, whereas $\tilde{\varphi}(\eta) = 1$. \triangleright

7.3.5. We continue presenting auxiliary propositions.

(1) *If Theorem 7.3.4 applies to each of the triples (G', G'_0, G'_f) and (G'', G''_0, G''_f) then so is $(G' \times G'', G'_0 \times G''_0, G'_f \times G''_f)$. In this event, $(G' \times G'')^\#$ is topologically isomorphic to $G'^\# \times G''^\#$.*

◁ The proof is easy and thus omitted. ▷

From 7.3.5 (1) it follows then in the case of $G_f = G$ under consideration that $(G \times G)^\# := (G \times G)/(G_0 \times G_0)$ is topologically isomorphic to $G^\# \times G^\# := (G/G_0) \times (G/G_0)$.

(2) Assume that $K : G^{\#2} \rightarrow \mathbb{C}$ is a continuous function and $K := \tilde{k}$, with $k : G \times G \rightarrow {}^*\mathbb{C}$ an S -continuous internal function. Given $g \in G$, define the function $K_{g^\#} : G^\# \rightarrow \mathbb{C}$ by the rule $K_{g^\#}(\cdot) := K(g^\#, \cdot)$, and also define the internal function $k_g : G \rightarrow {}^*\mathbb{C}$ by the rule $k_g(\cdot) := k(g, \cdot)$. Then k_g is S -continuous, and $K_{g^\#} = \tilde{k}_g$.

◁ Obvious. ▷

(3) If $f : G^\# \rightarrow \mathbb{C}$ is an even continuous function then there is an internal even S -continuous function φ satisfying $f = \tilde{\varphi}$. Moreover, if $K : G^{\#2} \rightarrow \mathbb{C}$ and $k : G^2 \rightarrow {}^*\mathbb{C}$ are defined by the rules $K(\xi, \eta) := f(\xi - \eta)$ and $k(g_1, g_2) := \varphi(g_1 - g_2)$ then $K = \tilde{k}$.

◁ Given $f = \tilde{\psi}$, with $\psi : G \rightarrow {}^*\mathbb{C}$ an internal S -continuous function, put $\varphi(g) := \frac{1}{2}(\psi(g) + \psi(-g))$. ▷

7.3.6. We now turn to interplay between integral equations on $G^\#$ and simultaneous linear algebraic equations on G .

First of all we recall that since $G^\#$ is a compact group; every Haar measure μ is finite and we may treat μ as a probability measure on assuming that $\mu(G^\#) = 1$. This measure is in correspondence to the uniform Loeb measure on G with weight $\Delta := |G|^{-1}$. It follows from the definition of a lifting of a measurable function $f : G^\# \rightarrow \mathbb{C}$ (see 7.2.5 and 7.2.6) that if $f = \tilde{\varphi}$ for some pointwise limited S -continuous function φ then φ is a lifting of f . Further, $\varphi \in \mathcal{L}_p(G)$ for every $p \in [1, \infty)$.

In case $\Delta = |G|^{-1}$ we simply write $\mathcal{L}_2(G)$ instead of $\mathcal{L}_{2,\Delta}(G)$. Distinguishing the canonical orthonormal basis $(e_h)_{h \in G}$ for $\mathcal{L}_2(G)$, with $e_h(g) := |G|^{1/2} \delta_{hg}$, we now turn to the equations of the form

$$(1) \quad \varphi(g) = \lambda |G|^{-1} \sum_{h \in G} k(g, h) \varphi(h),$$

where k is a pointwise limited internal symmetric function.

If this equation has a nonzero solution then λ is an *eigenvalue*. Each solution is referred to as an *eigenfunction* with eigenvalue λ . Thus, the eigenvalues of (1) are the reciprocals of the nonzero eigenvalues of the operator A given in the canonical orthonormal basis by the matrix $(a_{gh})_{g,h \in G}$ with $a_{gh} := |G|^{-1} k(g, h)$.

We also consider the corresponding integral equation on the group $G^\#$:

$$(2) \quad f(\xi) = \gamma \int_{G^\#} \tilde{k}(\xi, \eta) f(\eta) d\mu(\eta),$$

with γ a standard real.

7.3.7. Assume that $k : G^2 \rightarrow {}^*\mathbb{C}$ is a pointwise limited S -continuous internal function and $\varphi : G \rightarrow {}^*\mathbb{C}$ is a pointwise limited internal function. Define the internal function $\psi : G \rightarrow {}^*\mathbb{C}$ by the formula

$$\psi(g) := |G|^{-1} \sum_{h \in G} k(g, h) \varphi(h) \quad (g \in G).$$

Then ψ is pointwise limited and S -continuous. Moreover, if φ is an S -continuous function then

$$\tilde{\psi}(\xi) = \int_{G^\#} \tilde{\varphi}(\eta) \tilde{k}(\xi, \eta) d\mu(\eta) \quad (\xi \in G^\#).$$

\triangleleft The function ψ is pointwise limited since so are k and φ . By 7.3.5 (3) and the S -continuity of k we see that $k(g_1, h) - k(g_2, h) \approx 0$ for all $h \in G$ whenever $g_1 - g_2 \in G_0$. In other words, there is some infinitesimal $\alpha \approx 0$ satisfying $|k(g_1, h) - k(g_2, h)| \leq \alpha$ for all $h \in G$. Let $C > 0$ be a standard real, for which $|\varphi(h)| \leq C$ for all $h \in G$. Then $|\psi(g_1) - \psi(g_2)| \leq C\alpha \approx 0$, implying that ψ is S -continuous. The second assertion follows from Theorem 6.5.3. \triangleright

7.3.8. We now list a few useful properties of the eigenvalues of the equation 7.3.6 (1).

(1) The equation 7.3.6 (1) has no infinitesimal eigenvalues. If λ is a limited eigenvalue and R_λ is the eigenspace with eigenvalue λ then $\dim(R_\lambda) \in \mathbb{N}$, i.e., $\dim(R_\lambda)$ is a standard natural.

\triangleleft Since k is pointwise limited, ${}^\circ \sum_{g, h \in G} |a_{gh}|^2 < +\infty$. Hence, A meets the hypotheses of 6.1.11 and so the claim follows. \triangleright

Below we assume that the pointwise limited symmetric internal function k in 7.3.6 (1) is S -continuous.

(2) If λ is a limited eigenvalue then to each eigenfunction $\varphi \in R_\lambda$ there is an S -continuous pointwise limited function proportional to φ .

\triangleleft If $\varphi \in R_\lambda$ and $\varphi \neq 0$ then $\varphi_1 := \varphi / \max\{|\varphi(g)| : g \in G\}$ is bounded and $\varphi_1 \in R_\lambda$. We are done on applying 7.3.7. \triangleright

If $\varphi \in \mathcal{L}_2(G)$ then $\|\varphi\|^2 = |G|^{-1} \sum_{g \in G} |\varphi(g)|^2$ and $\|\varphi\|_\infty = \max\{|\varphi(g)| : g \in G\}$. Thus, $\|\varphi\| \leq \|\varphi\|_\infty$. Note nevertheless the following:

(3) If φ is an eigenfunction of the equation 7.3.6 (1) with a limited eigenvalue λ and $\|\varphi\| = 1$ then φ is S -continuous.

◁ Suppose that $\varphi_1 := C\varphi$ is an S -continuous eigenfunction of 7.3.6(1) satisfying $\|\varphi_1\|_\infty = 1$ (which is available by (2)). Then $\|\tilde{\varphi}\|_\infty = 1$, yielding $\int_{G^\#} |\tilde{\varphi}_1|^2 d\mu > 0$. However, φ_1 is a lifting of $\tilde{\varphi}_1$ and ${}^\circ\|\varphi_1\|^2 = \int_{G^\#} |\tilde{\varphi}_1|^2 d\mu$ by the pointwise limitedness of these functions and 7.2.6. Thus, $0 < {}^\circ\|\varphi_1\| < +\infty$. Note now that $\varphi_2 = \varphi_1/\|\varphi_1\|$ is an S -continuous and pointwise limited function, with $\|\varphi_2\| = 1$. Since $\varphi_2 = C_1\varphi$, $\|\varphi\| = 1$, and $C_1 > 0$; therefore, it follows that $C_1 = 1$ implying that $\varphi_2 = \varphi$. ▷

The claim of 7.3.8(3) holds clearly for every eigenfunction φ satisfying $0 < {}^\circ\|\varphi\| < +\infty$.

7.3.9. Let $f : G^\# \rightarrow \mathbb{C}$ be a continuous solution of the integral equation 7.3.6(2), with γ a standard number, $\gamma \neq 0$.

Then there are a standard natural n , some eigenvalues $\lambda_1, \dots, \lambda_n$ of 7.3.5(1) and pointwise limited S -continuous eigenfunctions $\varphi_1, \dots, \varphi_n$ such that $\lambda_i \approx \gamma$ and $\varphi_i \in R_{\lambda_i}$ for all $i := 1, \dots, n$ and f is a linear combination of $\tilde{\varphi}_1, \dots, \tilde{\varphi}_n$.

◁ By Theorem 6.5.3 (cf. 7.3.7) the above operator A with matrix $a_{gh} := |G|^{-1/2}k(g, h)$ in the canonical orthonormal basis of $\mathcal{L}_2(G)$ is a hyperapproximant to the integral operator \mathcal{A} with kernel $\tilde{k} : G^{\#2} \rightarrow \mathbb{C}$. This means that the following diagram commutes:

$$\begin{array}{ccc} L_2(G^\#) & \xrightarrow{\mathcal{A}} & L_2(G^\#) \\ j_2 \downarrow & & \downarrow j_2 \\ \mathcal{L}_2(G)^\# & \xrightarrow{A} & \mathcal{L}_2(G)^\# \end{array}$$

Recall that the mapping j_2 assigns to each function $f \in L_2(G^\#)$ the coset of an $\mathcal{S}_2(G)$ -lifting of it; i.e., the $\mathcal{S}_2(G)$ -lifting of $f \circ j$, where $j : G \rightarrow G^\#$ is the quotient homomorphism. The diagram makes it clear that $j_2(f)$ is an eigenvector of $A^\#$ with the eigenvalue γ^{-1} . By 6.1.10, there is some eigenvalue $\lambda^{-1} \approx \gamma^{-1}$ of A . By 7.3.8(3), every normalized eigenfunction φ of A with eigenvalue λ^{-1} is S -continuous ($0 < {}^\circ|\lambda^{-1}| < +\infty$). Moreover, $\tilde{\varphi}$ is an eigenfunction of \mathcal{A} , with eigenvalue γ^{-1} by 7.3.7.

Since \mathcal{A} is a compact operator; therefore, the eigenvalue γ^{-1} of \mathcal{A} has finite multiplicity. Hence, there is only a standardly finite number of eigenvalues of A infinitely close to γ^{-1} , each of standardly finite multiplicity. This means that the conditions of 6.1.12 are fulfilled.

Hence, to some standard n there are $\lambda_1, \dots, \lambda_n \approx \gamma^{-1}$ and pairwise orthonormal $\varphi_1, \dots, \varphi_n \in \mathcal{L}_2(G)$ such that $\varphi_k \in R_{\lambda_k^{-1}}$ for all $k \leq n$ and $j_2(f) = \sum_{k=1}^n C_k \varphi_k^\#$. Since φ_k is S -continuous, φ_k is an $\mathcal{S}_2(G)$ -lifting of $\tilde{\varphi}_k$. Hence $\varphi_k^\# = j_2(\tilde{\varphi}_k)$, and so

$$f = \sum_{k=1}^n C_k \tilde{\varphi}_k. \quad \triangleright$$

Closing this section, we address the problem of the general form of irreducible unitary representations of $G^\#$.

Let V be an internal Hilbert space. A unitary representation T of a group G in V (i.e., T is a homomorphism from G to the bounded endomorphism space $B(V)$, with $T(g)$ a unitary operator for every $g \in G$) is S -continuous provided that $\|T(g) - I_V\| \approx 0$ for all $g \in G_0$, with I_V standing as usual with the identity mapping in V . Such a mapping $T : G \rightarrow B(V)$ is a *hyperrepresentation* provided that if V is an internal hyperfinite-dimensional Hilbert space: in symbols, $\dim(V) = n \in {}^*\mathbb{N}$.

In the sequel we deal only with hyperrepresentations. If $\dim(V)$ is a standard hypernatural then each S -continuous irreducible unitary representation $T : G \rightarrow {}^*B(V)$ determines the continuous representation \tilde{T} of $G^\#$ by the rule $\tilde{T}(g^\#) := {}^\circ T(g)$. This \tilde{T} is a unitary representation too. Moreover, the character \varkappa of \tilde{T} may be written as $\tilde{\chi}$, with χ a character of T . Therefore, $\|\varkappa\| = {}^\circ\|\chi\| = 1$, since T is irreducible. This implies that \tilde{T} is irreducible too.

It is worth noting that the converse is also true. Precisely, we have the following abstraction of Theorem 7.2.11 (2) for commutative groups.

7.3.10. Theorem. *Each S -continuous irreducible unitary representation T of G generates an irreducible unitary representation \tilde{T} of $G^\#$ by the formula $\tilde{T}(g^\#) := {}^\circ T(g)$ for all $g \in G$. Conversely, each irreducible unitary representation of $G^\#$ has the shape \tilde{T} for some S -continuous irreducible unitary representation T of G .*

◁ To prove the first claim it suffices to show that each S -continuous irreducible unitary representation has finite rank. This will be done in 7.3.11.

The second claim follows from 7.3.12 by the well-known properties of irreducible unitary representations of a compact group (for example, see [377; Chapter 6, Section 32]). ▷

7.3.11. *Every S -continuous irreducible unitary representation of G has standardly finite rank.*

◁ Given $\xi \in V$, consider the sesquilinear form

$$\varphi_\xi(\eta, \zeta) := |G|^{-1} \sum_{g \in G} (T(g)\xi, \eta) \overline{(T(g)\xi, \zeta)}.$$

Let $B_\xi : V \rightarrow V$ be the linear operator acting by the rule $\varphi_\xi(\eta, \zeta) := (B_\xi\eta, \zeta)$.

Simple calculation shows that B_ξ commutes with every operator $T(g)$, and so by the Schur Lemma $B_\xi = \alpha(\xi)I$, with $\alpha(\xi) \in {}^*\mathbb{C}$.

Therefore, $\varphi_\xi(\eta, \zeta) = \alpha(\xi)(\eta, \zeta)$. Putting $\eta := \zeta$, find $\varphi_\xi(\zeta, \zeta) = \alpha(\xi)\|\zeta\|^2 = \alpha(\zeta)\|\xi\|^2$. The last equality implies that there is some hyperreal D in ${}^*\mathbb{R}$ such that $\alpha(\xi) = D\|\xi\|^2$ for all $\xi \in V$. Assume that the vector ξ is a norm-one vector. Then

$\varphi_\xi(\xi, \xi) = \alpha(\xi) = D$. Hence,

$$D = |G|^{-1} \sum_{g \in G} |(T(g)\xi, \xi)|^2$$

for every norm-one vector $\xi \in V$.

Show now that ${}^\circ D > 0$. To this end, consider the internal function $\psi : G \rightarrow {}^*\mathbb{R}$ given by $\psi(g) := |(T(g)\xi, \xi)|^2$. It is easy that ψ is S -continuous. Consequently, $\|\tilde{\psi}\|^2 = {}^\circ\|\psi\|^2 = D$, where the left side implies the $L_2(G^\#)$ norm. The definition of ψ yields $\psi(e) = 1$, where e is the identity of G . Consequently, $\tilde{\psi}(e^\#) = 1$. Since $\tilde{\psi}$ is a continuous function; therefore, $\|\psi\| > 0$ as required.

Now, let $\theta_1, \dots, \theta_n \in V$ form some orthonormal basis for V . Then

$$|G|^{-1} \sum_{g \in G} |(T(g)\theta_k, \theta_1)|^2 = \varphi_{\theta_k}(\theta_1, \theta_1) = \alpha(\theta_k)\|\theta_1\|^2 = D.$$

Since $T(g)$ is a unitary operator, the family $\{T(g)\theta_k : k := 1, \dots, n\}$ is an orthonormal basis; i.e.,

$$\sum_{k=1}^n |(T(g)\theta_k, \theta_1)|^2 = \|\theta_1\|^2 = 1.$$

Summing the last equality over g and multiplying by $|G|^{-1}$, from the previous equality we infer that $nD = 1$. The standardness of n is now clear since ${}^\circ D > 1$. \triangleright

7.3.12. *The linear span of all functions $\tilde{\psi}$, with $\psi(g)$ a matrix entry of some S -continuous irreducible unitary representation of G , is dense in $C(G^\#)$.*

\triangleleft Inspection of the proof of Theorem 32 of [402] shows that the linear span of eigenfunctions of all integral equations with kernels like $f(x-y)$, with $f : G^\# \rightarrow \mathbb{C}$ a continuous even function, is dense in $C(G^\#)$. By 7.3.5 (3) and 7.3.9, the linear space of linear combinations of functions like $\tilde{\varphi}$, with φ an S -continuous eigenfunction of some equation of the form 7.3.6 (1) with $0 < {}^\circ|\lambda| < +\infty$ and $k(g, h) := f(g-h)$ for an appropriate S -continuous even function $f : G \rightarrow {}^*\mathbb{C}$, is dense in $C(G^\#)$.

The rest of the proof proceeds in fact along the lines of the proof of Theorem 32 in [402] and is presented here only for the sake of completeness.

We so assume that $k(g, h) := f(g-h)$ in 7.3.5 (1) for an even continuous function $f : G^\# \rightarrow \mathbb{C}$. By 7.3.8 (1), the dimension of R_λ is standardly finite. Let $\varphi_1, \dots, \varphi_n$ form a complete orthonormal system of eigenfunctions of 7.3.6 (1). Moreover, we may and will presume the latter S -continuous. Obviously, if $\varphi(g) \in R_\lambda$, then $\varphi(a+g) \in R_\lambda$ for all $a \in G$. Thus, $\varphi_1(a+g), \dots, \varphi_n(a+g)$ also comprise a complete orthonormal system of eigenfunctions of 7.3.6 (1), and so there is a unitary

matrix $U(a) := (u_{\iota_j}(a))_{\iota,j=1}^n$ satisfying

$$\varphi_\iota(a + g) = \sum_{j=1}^n u_{\iota_j}(a)\varphi_j(g).$$

Show now that $\{U(a) : a \in G\}$ is a representation of G . Indeed, $u_{\iota_j}(a + b) = \sum_{k=1}^n u_{\iota_k}(a)u_{k_j}(b)$. Since $\{\varphi_\iota : \iota := 1, \dots, n\}$ is orthonormal,

$$u_{\iota_j}(a) = |G|^{-1} \sum_{g \in G} \varphi_\iota(a + g)\varphi_j(g).$$

Since φ_ι are pointwise limited and S -continuous, the last equality implies that $u_{\iota_j}(a)$ is S -continuous. Since $U(\cdot)$ is a unitary representation of G , there is a unitary matrix V such that $U(a) = VX(a)V^{-1}$ for all $a \in G$, with

$$X(a) := \begin{pmatrix} T_1(a) & & 0 \\ & \ddots & \\ 0 & & T_n(a) \end{pmatrix},$$

and T_ι is an irreducible unitary representation of G for all $\iota := 1, \dots, n$.

Since $X(a) = V^{-1}U(a)V$, all matrix entries of each representation T_ι are standardly finite linear combinations of S -continuous functions, that is, the representations T_ι are themselves S -continuous. Similarly, the entries u_{ι_j} are standardly finite linear combinations of matrix entries of T_ι 's with limited coefficients.

Putting $g := 0$ in the above expression for $\varphi_\iota(a + g)$ infer that each φ_ι is a standardly finite linear combination of u_{ι_j} 's with limited coefficients, and hence also of some matrix entries ψ_k of the representations T_ι 's. Clearly, if $\varphi_\iota = \sum_{j=1}^n C_j\psi_j$ then $\tilde{\varphi}_\iota = \sum_{j=1}^n \circ(C_j)\tilde{\psi}_j$. \triangleright

7.3.13. COMMENTS.

(1) By Theorem 7.2.3, $G^\#$ is a compact group if and only if G_f is an internal subgroup of G . We may thus assume without loss of generality that $G_f = G$. By 7.2.11 it suffices to show that each character $\varkappa \in G^{\#\wedge}$ has the shape $\tilde{\chi}$ with some $\chi \in G^\wedge$ satisfying the condition $\chi|_{G_0} \approx 1$ (cf. 7.2.7 (3)). The last is easy since the system of all character like $\tilde{\chi}$ is complete. This completeness ensues in turn from a slight modification of the Peter–Weyl Theorem asserting the completeness of the character set of irreducible representations of a compact group (cf. [402]).

(2) It is worth noting that all considerations of this section but the results on the character group remain valid on assuming that G is internal hyperfinite nonabelian group, while the external subgroups G_0 and G_f , with $G_0 \subset G_f$, satisfying the conditions (A) and (B) of 7.2.1, are normal subgroups of G .

(3) The proof of 7.3.11 is analogous to that of Theorem 22.13 in [174] asserting that each irreducible representation of a compact group has finite rank. However the context of 7.3.11 is easier since we deal here with hyperfinite groups which may be treated as finite groups in many aspects.

7.4. Hyperapproximation of Locally Compact Abelian Groups

In this section we turn to hyperapproximation of a topological group which is in fact the main topic of Chapter 7. Most results concern the case of locally compact abelian group.

7.4.1. Recall that if \mathfrak{G} is a topological abelian group, then the *monad* $\mu_{\mathfrak{G}}(0)$ of G (i.e. the monad of the neighborhood filter of zero) and the *nearstandard part* $\text{nst}(*\mathfrak{G})$ of G are determined by the formulas

$$\begin{aligned}\mu_{\mathfrak{G}}(0) &:= \bigcap \{ {}^*U : 0 \in U, U \subset \mathfrak{G}, U \text{ is open} \}, \\ \text{nst}(*\mathfrak{G}) &:= \{ \xi \in *\mathfrak{G} : (\exists \eta \in \mathfrak{G})(\xi \approx \eta) \},\end{aligned}$$

with $\xi_1 \overset{\mathfrak{G}}{\approx} \xi_2 := \xi_1 \approx \xi_2$ meaning that $\xi_1 - \xi_2 \in \mu_{\mathfrak{G}}(0)$.

The mapping $\text{st} : \text{nst}(*\mathfrak{G}) \rightarrow \mathfrak{G}$, acting by the rule $\text{st}(\xi) \approx \xi$ for all $\xi \in \text{nst}(*\mathfrak{G})$, is an epimorphism with kernel $\mu_{\mathfrak{G}}(0)$, and so, $\mathfrak{G} \cong \text{nst}(*\mathfrak{G})/\mu_{\mathfrak{G}}(0)$. We simply write $\mu(0)$ and $\xi_1 \approx \xi_2$ instead of $\mu_{\mathfrak{G}}(0)$ and $\xi_1 \overset{\mathfrak{G}}{\approx} \xi_2$ when this leads to no confusion. Now the main definition is in order.

Assume that \mathfrak{G} is a standard topological group, G is an internal hyperfinite group, and $j : G \rightarrow *\mathfrak{G}$ is an internal mapping. The couple (G, j) is a *hyperapproximant* of \mathfrak{G} provided that the following hold:

- (1) To each $\xi \in \text{nst}(*\mathfrak{G})$ there is some g in G satisfying $j(g) \approx \xi$;
- (2) If $g_1, g_2 \in j^{-1}(\text{nst}(*\mathfrak{G}))$ then $j(g_1 + g_2) \approx j(g_1) + j(g_2)$;
- (3) If $g \in j^{-1}(\text{nst}(*\mathfrak{G}))$ then $j(-g) \approx -j(g)$;
- (4) $j(0) = 0$.

We could have postulated that $j(0) \approx 0$ in (4) of course, but we may the exact equality by an obvious change while keeping (1)–(3) intact.

It is worth noting that the above definition does not presume that the group G under study nor any hyperapproximant to it are abelian, whereas we still use the symbols $+$ and 0 for denoting the group operation and the neutral element of G . However, if G is abelian then a hyperapproximant of G is also presumed abelian.

Put $G_f := j^{-1}(\text{nst}(*\mathfrak{G}))$, $G_0 := j^{-1}(\mu(0))$, and $\tilde{j} := \text{st} \circ j|_{G_f}$. Then the conditions (1)–(4) amount to requiring that $\tilde{j} : G_f \rightarrow \mathfrak{G}$ be an epimorphism with kernel $\ker(\tilde{j}) = G_0$. We let \bar{j} stand for the induced isomorphism \bar{j} between $G^{\#} := G_f/G_0$ and \mathfrak{G} . The quotient homomorphism of G_f onto $G^{\#}$ is denoted simply by $\#$.

7.4.2. Say that (G, j) is a *nice hyperapproximant* to a separable locally compact abelian group \mathfrak{G} provided that the corresponding triple (G, G_0, G_f) is admissible (cf. 7.2.13).

As a helpful example of nice hyperapproximation, we consider the additive group $G := \{-L, \dots, L\}$ of the ring ${}^*\mathbb{Z}/N{}^*\mathbb{Z}$, where $N := 2L + 1$ is an infinite hypernatural and $\Delta \approx 0$ is a strictly positive infinitesimal satisfying $N\Delta \approx +\infty$. Define the mapping $j : G \rightarrow {}^*\mathbb{R}$ by putting $j(k) := k\Delta$ for all $k \in G$. Obviously, (G, j) is a hyperapproximant to the additive group of the reals \mathbb{R} . Theorem 7.1.2, together with 7.2.1 (1), shows that this (G, j) is also a nice hyperfinite approximant.

7.4.3. *If \mathfrak{G} is a separable locally compact group and (G, j) is its hyperapproximant then the triple (G, G_0, G_f) satisfies the conditions of Theorem 7.2.3. Moreover, $\bar{j} : G^\# \rightarrow \mathfrak{G}$ is a topological isomorphism.*

◁ Since \mathfrak{G} is locally compact and separable, assume that $\mathfrak{G} = \bigcup_{n=1}^\infty U_n$ where each U_n is a relatively compact open set. It is then easy that $G_f = \bigcup_{n=1}^\infty j^{-1}(*U_n)$. By exact analogy, if $\{V_n : n \in \mathbb{N}\}$ is a countable neighborhood base of zero of \mathfrak{G} consisting of relatively compact sets then $G_0 = \bigcap_{n \in \mathbb{N}} j^{-1}(*V_n)$. Therefore, G_f is a countable union and G_0 is a countable intersection of internal sets, which implies that $G^\#$ admits the canonical topology by Theorem 7.2.2. The hypotheses of Theorem 7.2.3 follow on observing that $G^\#$ is a locally compact group. We are left with proving that \bar{j} and \bar{j}^{-1} are continuous at zero.

Given a neighborhood V about the zero of \mathfrak{G} , find a relatively compact neighborhood about the zero of V^1 so that $V^1 \subset V$. Consider the internal set $F := j^{-1}(*V^1)$. Clearly, $G_0 \subset F$ and so $\overset{\circ}{F}^\#$ is a neighborhood about the zero of $G^\#$. The continuity of \bar{j} at zero follows from the easy relation $\bar{j}(\overset{\circ}{F}^\#) \subset V^1 \subset V$.

Take a neighborhood about the zero of $G^\#$ of the shape $\overset{\circ}{F}^\#$, with F an internal subset of G_f including G_0 . Since $G_0 = \bigcap \{j^{-1}(*U) : U \text{ is a neighborhood about the zero of } \mathfrak{G}\}$, by ω^+ -saturation we infer that there is a relatively standard neighborhood U about the zero of \mathfrak{G} satisfying $j^{-1}(*U) \subset F$. Let V be a neighborhood about the zero of \mathfrak{G} such that $V + V + V \subset U$. From the obvious inclusion $V \subset V + V$, it follows that $j^{-1}(*V) + j^{-1}(*V) \subset j^{-1}(*U)$. Therefore, if $\xi \in V$ and $\bar{j}^{-1}(\xi) = g^\#$ or, which is the same, $j(g) \approx \xi$ then $g \in \overset{\circ}{F}$. Hence, $\bar{j}^{-1}(V) \subset \overset{\circ}{F}^\#$. ▷

7.4.4. Theorem. *Each separable locally compact abelian group with a compact and open subgroup possesses a nice hyperapproximant.*

◁ Theorems 7.2.12 and 7.2.14 show that in this case every hyperapproximant is nice, and so it suffices to prove the existence of some hyperapproximant.

(1) Assume that G is a separable locally compact abelian group with a compact and open subgroup U . Denote the quotient group \mathfrak{G}/U by \mathscr{D} and consider

the short exact sequence $U \subset \mathfrak{G} \xrightarrow{\pi} \mathcal{D}$, with π the quotient homomorphism. By ω^+ -saturation and the countability of \mathcal{D} there is a hyperfinite set $T \subset {}^*\mathcal{D}$ satisfying $\mathcal{D} \subset T$. Let ${}^*\mathcal{D}(T)$ stand for the internal subgroup of ${}^*\mathcal{D}$ generated by T and put $\mathcal{H} := {}^*\pi^{-1}({}^*\mathcal{D}(T))$. Thus, the following diagram commutes

$$\begin{array}{ccccc} U & \xrightarrow{\subset} & \mathfrak{G} & \xrightarrow{\pi} & \mathcal{D} \\ \text{id} \downarrow & & \text{id} \downarrow & & \text{id} \downarrow \\ U & \xrightarrow{\subset} & \mathcal{H} & \xrightarrow{\varepsilon} & {}^*\mathcal{D}(T), \end{array}$$

with the bottom row a short exact sequence and $\varepsilon := {}^*\pi|_{{}^*\mathcal{H}}$ an internal mapping.

Recall that a finitely generated abelian group splits in the direct sum of a free subgroup and a finite subgroup (see [305; §10, Theorem 8]). By transfer, infer that ${}^*\mathcal{D}(T) = \mathcal{D}_1 \oplus \mathcal{D}_2$, where \mathcal{D}_1 is a hyperfinite abelian group and \mathcal{D}_2 is a (non-standardly) free abelian hyperfinitely generated group. Putting $\mathcal{H}_i := \varepsilon^{-1}(\mathcal{D}_i)$ for $i := 1, 2$, note that $\mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$ and $\mathcal{H}_1 \cap \mathcal{H}_2 = {}^*U$ while the following are exact sequences:

$${}^*U \subset \mathcal{H}_1 \xrightarrow{\varepsilon_1} \mathcal{D}_1; \quad {}^*U \subset \mathcal{H}_2 \xrightarrow{\varepsilon_2} \mathcal{D}_2,$$

with $\varepsilon_i := \varepsilon|_{\mathcal{H}_i}$ for $i := 1, 2$.

Look at these sequences thoroughly. Start with the first of them, applying transfer to the van Kampen Theorem (see [174, Chapter 2, Theorem 9.5]). Given an infinitesimal neighborhood V of zero (which means $V \subset \mu(0)$) with $V \subset {}^*U$, find a hypernatural k , a hyperfinite group R , and a continuous epimorphism $\varphi : {}^*U \rightarrow {}^*S^k \oplus R$ satisfying $\ker(\varphi) \subset V$. Here S is the unit circle.

(2) Let R be a normal subgroup of a group L with the quotient group L/R isomorphic H . In this event L is an *extension* of R by H ; in symbols, $\text{Ext}(H, R) := L$.

In these terms, the above means that \mathcal{H}_1 is an extension of *U by \mathcal{D}_1 . Since φ is an epimorphism, there is an extension L of ${}^*S^k \oplus R$ by \mathcal{D}_1 , so that for some internal group L and an internal homomorphism $\gamma : \mathcal{H}_1 \rightarrow L$ the following diagram commutes:

$$\begin{array}{ccccc} {}^*U & \xrightarrow{\subset} & \mathcal{H}_1 & \xrightarrow{\varepsilon_1} & \mathcal{D}_1 \\ \varphi \downarrow & & \downarrow \gamma & & \downarrow \text{id} \\ {}^*S^k \oplus R & \xrightarrow{\simeq} & L & \xrightarrow{\delta} & \mathcal{D}_1, \end{array}$$

with the bottom row a short exact sequence. This commutative diagram implies that γ is an epimorphism.

The easy equality $\text{Ext}(\mathcal{D}_1, S^k \oplus R) = \text{Ext}(\mathcal{D}_1, S^k) \oplus \text{Ext}(\mathcal{D}_1, R)$ implies the existence of the two short exact sequences

$$\begin{aligned} {}^*S^k &\xrightarrow{\varkappa_1} L_1 \xrightarrow{\delta_1} \mathcal{D}_1, \\ R &\xrightarrow{\varkappa_2} L_2 \xrightarrow{\delta_2} \mathcal{D}_1, \end{aligned}$$

yielding the bottom row of the above diagram to within isomorphism as follows: $L := \{(l_1, l_2) \in L_1 \oplus L_2 : \delta_1(l_1) = \delta_2(l_2)\}$, $\varkappa := (\varkappa_1, \varkappa_2)$, $\delta(l_1, l_2) := \delta_1(l_1) = \delta_2(l_2)$. Since ${}^*S^k$ is a divisible group, the first of these sequences splits; i.e., there is a monomorphism $\chi : \mathcal{D}_1 \rightarrow L_1$ serving as a right inverse of δ_1 . The group L_2 is hyperfinite since so is each of the groups R and \mathcal{D}_1 .

Note that φ is open as a continuous epimorphism of compact groups, and so $\varphi(V)$ is a neighborhood about the zero of ${}^*S^k \oplus R$, implying that $\varphi(V) \cap {}^*S^k$ is a neighborhood about the zero of ${}^*S^k$. Given an arbitrarily small $\varepsilon > 0$, we may find in the unit circle S a finite subgroup that is an ε -net for S . Consequently, there is a hyperfinite subgroup $F \subset {}^*S^k$ satisfying $F + (\varphi(V) \cap {}^*S^k) = {}^*S^k$. We now consider the hyperfinite subgroup $M \subset L$ given by

$$M := \{(\varkappa_1(f) + \chi(d), l) : f \in F, d \in \mathcal{D}_1, l \in L_2, \delta_2(l) = d\}.$$

Since γ is an epimorphism; therefore, $\gamma^{-1}(m) \neq \emptyset$ for all $m \in M$. Choose a member g_m from each set $\gamma^{-1}(m)$ by internal choice, and put $G_1 := \{g_m : m \in M\}$. Define the operation $+_1$ in G_1 by the rule $g_{m_1} +_1 g_{m_2} := g_{m_1+m_2}$. Thus, $\gamma(g_{m_1} +_1 g_{m_2}) = \gamma(g_{m_1}) + \gamma(g_{m_2}) = m_1 + m_2$. Obviously, $(G_1, +_1)$ is a hyperfinite abelian group. We will need a few properties of $(G_1, +_1)$.

(3) *If $m, m_1, m_2 \in M$ then $g_{m_1} +_1 g_{m_2} \approx g_{m_1} + g_{m_2}$ and $-_1 g_m \approx -g_m$.*

\triangleleft Since $\gamma(g_{m_1} +_1 g_{m_2}) = \gamma(g_{m_1}) + \gamma(g_{m_2})$, the commutative diagram (2) yields $\varepsilon_1(g_{m_1} +_1 g_{m_2} - g_{m_1} - g_{m_2}) = 0$, i.e., $g_{m_1} +_1 g_{m_2} - g_{m_1} - g_{m_2} \in {}^*U$. Using the left square of this diagram and the fact that \varkappa is a monomorphism, infer that $\varphi(g_{m_1} +_1 g_{m_2} - g_{m_1} - g_{m_2}) = 0$. The last formula may be rewritten as $g_{m_1} +_1 g_{m_2} - g_{m_1} - g_{m_2} \in V \subset \mu(0)$, which proves the first claim. The proof of the claim assertion proceeds along the same lines. \triangleright

(4) *To each $h \in \mathcal{H}_1$ there is some g in G_1 satisfying $g \approx h$.*

\triangleleft Take an arbitrary element $h \in \mathcal{H}_1$. Using the structure of the group L and the splitting of (2) ${}^*S^k \xrightarrow{\varkappa_1} L_1 \xrightarrow{\delta_1} \mathcal{D}_1$, find that $\gamma(h) = (\varkappa_1(s) + \chi(d), l)$, where $l \in L_2$ and $\delta_2(l) = d$. Since $F + (\varphi(V) \cap {}^*S^k) = {}^*S^k$, there is some f in F satisfying $s - f \in \varphi(V) \cap {}^*S^k$. But then $m = (\varkappa_1(f) + \chi(d), l) \in M$. Show now that $g_m \approx h$. To this end note that $\varepsilon_1(h) = \varepsilon_1(g_m) = d$, and so $l - g_m \in {}^*U$. Moreover, $\gamma(l) - \gamma(g_m) = (\varkappa_1(f - s), 0) = \varkappa((f - s), 0) = \varkappa\varphi(l - g_m)$. Thus, $\varphi(l - g_m) = (f - s, 0) \in \varphi(V) \cap {}^*S^k \subset \varphi(V)$, implying that $l - g_m \in \varphi^{-1}(\varphi(V)) = V + \ker(\varphi) \subset V + V \subset \mu(0)$. \triangleright

(5) The group $G_U := {}^*U \cap G_1$ is a subgroup of G_1 . The couple (G_U, \hookrightarrow) , with \hookrightarrow the identity embedding of G_U into *U , is a hyperapproximant to U .

◁ This is immediate from (3) and (4) since U is a compact and open subgroup of \mathfrak{G} . ▷

(6) We turn now to inspecting the exact sequence ${}^*U \hookrightarrow \mathcal{H}_2 \xrightarrow{\varepsilon_2} \mathcal{D}_2$; cf. (1).

Let $\nu_i : {}^*\mathcal{D}(T) \rightarrow \mathcal{D}_i$ be the quotient mapping. Choose a hypernatural m such that $(\nu_2(T) - \nu_2(T)) \cap m\mathcal{D}_2 = 0$. The existence of m is easy by transfer since if P is a finite subset of a free finitely generated abelian group H then there is a natural m satisfying $P \cap mH \subset 0$.

Put $Q := \mathcal{D}_2/m\mathcal{D}_2$. Then Q is a hyperfinite abelian group. Denote the quotient homomorphism from \mathcal{D}_2 onto Q by λ . By construction, $\nu_2(T)$ is injective. By internal choice, take an element d_q in each $\lambda^{-1}(q)$, with $q \in Q$, so that if $\lambda^{-1}(q) \cap \nu_2(T) \neq \emptyset$ then $d_q \in \nu_2(T)$ (this d_q is unique because λ is a monomorphism on $\nu_2(T)$). Put $G_3 := \{d_q : q \in Q\}$. Define some operation $+_3$ over G_3 by the rule $d_{q_1} +_3 d_{q_2} := d_{q_1+q_2}$. Thus, $\lambda(d_{q_1} +_3 d_{q_2}) = \lambda(d_{q_1}) + \lambda(d_{q_2})$.

(7) To each $d \in \mathcal{D}$ there is some q in Q satisfying $\nu_2(d) = d_q$. If $d_{q_1}, d_{q_2} \in \nu_2(\mathcal{D})$ then $d_{q_1} +_3 d_{q_2} = d_{q_1} + d_{q_2}$.

Note that $+$ on the right side of the last equality stands for addition in \mathcal{D}_2 while $\nu_2(\mathcal{D})$ is an external subgroup of \mathcal{D}_2 in general.

◁ The first claim follows from the definition of d_q since $\nu_2(\mathcal{D}) \subset \nu_2(T)$. If $d_{q_1}, d_{q_2} \in \nu_2(\mathcal{D})$ then $d_{q_1} + d_{q_2} \in \nu_2(\mathcal{D})$. Put $d_{q_1} + d_{q_2} := d_{q_3}$. Since λ is a homomorphism; therefore, $q_3 = \lambda(d_{q_1} + d_{q_2}) = \lambda(d_{q_1}) + \lambda(d_{q_2}) = q_1 + q_2$. On the other hand, $\lambda(d_{q_1} +_3 d_{q_2}) = \lambda(d_{q_1}) + \lambda(d_{q_2}) = q_1 + q_2 = q_3$. Thus, $\lambda^{-1}(q_3) \cap \nu_2(T) = \{d_{q_1} + d_{q_2}\}$, and so $d_{q_3} = d_{q_1} + d_{q_2} = d_{q_1} +_3 d_{q_2}$. ▷

(8) Denote the restriction of the group operation $+_1$ to G_U by $+_U$. Since \mathcal{D}_2 is a free abelian group, the sequence ${}^*U \subset \mathcal{H}_2 \xrightarrow{\varepsilon_2} \mathcal{D}_2$ splits; i.e. there is a monomorphism $\mu_2 : \mathcal{D}_2 \rightarrow \mathcal{H}_2$ which is a right inverse of ε_2 . Consider the set $G_2 := \{g + \mu(d_q) : g \in G_U, q \in Q\}$ and define the operation $+_2$ on G_2 by putting $(g_1 + \mu(d_{q_1})) +_2 (g_2 + \mu(d_{q_2})) := g_1 +_U g_2 + \mu(d_{q_1} +_3 d_{q_2})$. Since $\mathcal{H}_2 = {}^*U \oplus \mu_2(\mathcal{D}_2)$ and $G_U \subset {}^*U$, it is easy to see that $(G_2, +_2)$ is a hyperfinite abelian group and $G_2 \cap {}^*U = G_U$.

(9) Suppose now that $\mathfrak{G} = G_1 \times G_2$. Define $j : G \rightarrow {}^*\mathfrak{G}$ by the rule $j((g_1, g_2)) := g_1 + g_2$. Then (G, j) is a hyperapproximant to \mathfrak{G} .

◁ Take $\xi \in \mathfrak{G}$. Since $\mathfrak{G} \subset \mathcal{H} = \mathcal{H}_1 + \mathcal{H}_2$ (cf. (1)); therefore, $\xi = h_1 + h_2$, with $h_1 \in \mathcal{H}_1$ and $h_2 \in \mathcal{H}_2$. By the diagram of (1), $d := \pi(\xi) = \varepsilon_1(h_1) + \varepsilon_2(h_2)$. Thus, $\varepsilon_2(h_2) = \nu_2(d) \in \nu_2(\mathcal{D})$ because $d \in \mathcal{D}$. By (7), $\nu_2(d) = d_q$. Hence, there is some g in *U satisfying $h_2 = g + \mu(d_q)$. By (5), G_U approximates *U , and since U is compact, there is some g_0 in G_U such that $g_0 \approx g$. Hence, $g_2 = g_0 + \mu(d_q) \approx$

$g + \mu(d_q) = h_2$ and $g_2 \in G_2$. By (3) there is some g_1 in G_1 satisfying $g_1 \approx h_1$. Consequently, $g_1 + g_2 \approx h_1 + h_2 = \xi$. This ensures the first condition of the definition of hyperapproximant (cf. 7.4.1 (1)). Since the fourth condition is obvious, it remains to verify 7.4.1 (2) and 7.4.1 (3).

Suppose that $g_1 + g_2 \approx \xi \in \mathfrak{G}$ and $g'_1 + g'_2 \approx \xi' \in \mathfrak{G}$, with $g_1, g'_1 \in G_1$ and $g_2, g'_2 \in G_2$. The claim is that $g_1 +_1 g'_1 + g_2 +_2 g'_2 \approx \xi + \xi'$. By (3), $g_1 +_1 g'_1 \approx g_1 + g'_1$. Consequently, it suffices to show that $g_2 +_2 g'_2 \approx g_2 + g'_2$. To this end, put $d := \pi(\xi)$. Since $g_1 + g_2 \approx \xi$; therefore, $g_1 + g_2 - \xi \in {}^*U$ and so $\varepsilon(g_1 + g_2) = d$ (see the diagram of (1)). From $g_i \in \mathcal{H}_i$ it follows that $\varepsilon_2(g_2) = \nu_2(d) \in \nu_2(\mathcal{D})$. Similarly, $\varepsilon_2(g'_2) = \nu_2(d'_2) \in \nu_2(\mathcal{D})$, with $d' := \pi(\xi')$. This implies easily that $g_2 = \tilde{g} + \mu(\nu_2(d))$ and $g'_2 = \tilde{g}' + \mu(\nu_2(d'))$. By (7) note now that $g_2 +_2 g'_2 = \tilde{g} +_U \tilde{g}' + \mu(\nu_2(d) + \nu_2(d')) = \tilde{g} +_U \tilde{g}' + \mu(\nu_2(d)) + \mu(\nu_2(d'))$. Also, by (5) $\tilde{g} +_U \tilde{g}' \approx \tilde{g} + \tilde{g}'$, which yields 7.4.1 (2). Check 7.4.1 (3) by analogy. \triangleright

We have thus completed the proof of Theorem 7.4.4. \triangleright

7.4.5. Theorem. *Each separable locally compact abelian group possesses a nice hyperapproximant.*

\triangleleft Each locally compact abelian group is presented as the product of \mathbb{R}^m for some $m \geq 0$ and a group with a compact and open subgroup (for example, see [151; Chapter 2, §10.3, Theorem 1]). It is also clear that if \mathfrak{G}_1 and \mathfrak{G}_2 are separable locally compact abelian groups with nice hyperapproximants then $\mathfrak{G}_1 \times \mathfrak{G}_2$ also possesses a nice hyperapproximant. We are done on appealing to Theorem 7.4.4 and recalling 7.4.2. \triangleright

7.4.6. Let (G, j) be a hyperapproximant to a group \mathfrak{G} . If $U \subset \mathfrak{G}$ is a relatively compact neighborhood about zero then $G_0 \subset j^{-1}({}^*U) \subset G_f$, implying that $\Delta := |j^{-1}({}^*U)|^{-1}$ is a normalizing factor of the triple (G, G_0, G_f) (cf. 7.2.4 (4)). The external subgroups G_0 and G_f , determined by the hyperapproximant (G, j) of \mathfrak{G} , lead in turn to the external subgroups $H_0, H_f \subset \widehat{G}$ (cf. 7.2.7).

If (G, j) is a hyperapproximant to \mathfrak{G} then it is natural to approximate the dual $\widehat{\mathfrak{G}}$ of \mathfrak{G} by \widehat{G} . If (\widehat{G}, ι) is a hyperapproximant to $\widehat{\mathfrak{G}}$ then $\widehat{G}_f = \iota^{-1}(\text{nst}({}^*\widehat{\mathfrak{G}}))$ and $\widehat{G}_0 = \iota^{-1}\mu_{\widehat{\mathfrak{G}}}(\mathbf{1})$. Here $\mathbf{1}$ is the neutral element of $\widehat{\mathfrak{G}}$, i.e., the identically one character on \mathfrak{G} .

Let (G, j) and (\widehat{G}, ι) be hyperapproximants to the separable locally compact abelian groups \mathfrak{G} and $\widehat{\mathfrak{G}}$, respectively. The couple (\widehat{G}, ι) is a *dual* of (G, j) provided that the following hold:

- (1) $H_0 \subset \widehat{G}_0$;
- (2) $\iota(h)(j(g)) \approx h(g)$ for all $h \in \widehat{G}_f$ and $g \in G_f$.

Note that if \mathfrak{G} is a compact group then $G_f = G$ and (1) holds automatically by 7.2.11 (1).

The Loeb measure ν_Δ on G induces the Haar measure μ_Δ on $G^\#$. The topological isomorphism \bar{j} transforms μ_Δ into the Haar measure $\bar{\mu}_\Delta$ on \mathfrak{G} . Obviously, every Haar measure on \mathfrak{G} is obtainable in this manner.

If $f : \mathfrak{G} \rightarrow \mathbb{R}$ is a measurable function then each lifting of $f \circ \bar{j}$ is shortly called a *lifting* of f (cf. 7.2.6).

7.4.7. *Let (G, j) be a hyperapproximant to a separable locally compact abelian group \mathfrak{G} and Δ a normalizing factor of the triple (G, G_0, G_f) . Assume that $p \in [1, \infty)$ is standard and $f : \mathfrak{G} \rightarrow \mathbb{C}$. Then $f \in L_p(\mathfrak{G})$ if and only if f has an $\mathcal{S}_{p,\Delta}$ -integrable lifting. Further, if $p = 1$ and $\varphi : G \rightarrow {}^*\mathbb{C}$ is an $\mathcal{S}_{1,\Delta}$ -integrable lifting of f then*

$$\int f d\bar{\mu}_\Delta = \overset{\circ}{\left(\Delta \sum_{g \in G} \varphi(g) \right)}.$$

◁ This is an easy rephrasing of 7.2.6. ▷

It now follows directly from 6.4.9.(1) that the triple (G, j, Δ) is a hyperapproximant to the measure space $(\mathfrak{G}, \mu_\Delta)$ and Proposition 6.4.10 reads:

7.4.8. *If in the context of 7.4.7 $f \in L_1(\mathfrak{G})$ is a bounded Haar almost everywhere continuous function satisfying the condition*

$$(\forall B \in {}^*\mathcal{P}(G)) \left(B \subset G - G_f \rightarrow \Delta \sum_{g \in G} |{}^*f(j(g))| \approx 0 \right),$$

then $\varphi := {}^*f \circ j$ is an $\mathcal{S}_{1,\Delta}$ -integrable lifting of f and

$$\int f d\mu_\Delta = \overset{\circ}{\left(\Delta \sum_{g \in G} {}^*f(j(g)) \right)}.$$

7.4.9. *Assume that $\chi \in {}^*\widehat{\mathfrak{G}}$ and $\varkappa \in \widehat{\mathfrak{G}}$. Then $\chi \overset{\widehat{\mathfrak{G}}}{\approx} \varkappa$ if and only if $\chi(\xi) \approx \varkappa(\xi)$ for all $\xi \in \text{nst}({}^*\mathfrak{G})$.*

◁ Suppose that $\chi \approx {}^*\varkappa$. If $\xi \in \text{nst}({}^*\mathfrak{G})$ then $\xi \approx \eta \in \mathfrak{G}$. If u is a relatively compact neighborhood of η in \mathfrak{G} then $\xi \in \bar{u}$. By hypothesis, $(\chi - {}^*\varkappa \in {}^*\mathcal{W}(\bar{u}, \Lambda_k))$ for every standard k (see the proof of Theorem 7.2.8), which means that $\chi(\xi) \approx {}^*\varkappa(\xi)$.

Conversely, suppose that $\chi(\xi) \approx {}^*\varkappa(\xi)$ for all $\xi \in \text{nst}({}^*\mathfrak{G})$. In this event if F is a compact subset of \mathfrak{G} , then ${}^*F \subset \text{nst}({}^*\mathfrak{G})$. Hence, $(\chi(\xi) \approx {}^*\varkappa(\xi))$ for all $\xi \in {}^*F$, and so $\chi - {}^*\varkappa \in {}^*\mathcal{W}(F, \Lambda_k)$ for every standard k . ▷

7.4.10. Theorem. Assume that (G, j) is a nice hyperapproximant to a separable locally compact abelian \mathfrak{G} and (\widehat{G}, ι) is a dual hyperapproximant to $\widehat{\mathfrak{G}}$. Assume further that Δ is a normalizing factor of the triple (G, G_0, G_f) constructed from (G, j) . Then the following hold:

- (1) $(|G| \cdot \Delta)^{-1}$ is a normalizing factor of the triple $(\widehat{G}, \widehat{G}_0, \widehat{G}_f)$ constructed from (\widehat{G}, ι) ;
- (2) If $\mathcal{F} : L_2(\mathfrak{G}, \bar{\mu}_\Delta) \rightarrow L_2(\widehat{\mathfrak{G}}, \bar{\mu}_{\widehat{\Delta}})$ is the Fourier transform then \mathcal{F} preserves the inner product;
- (3) The discrete Fourier transform $\Phi_\Delta^G : \mathcal{L}_{2,\Delta}(G) \rightarrow \mathcal{L}_{2,\widehat{\Delta}}(\widehat{G})$ is a hyperapproximant to \mathcal{F} .

◁ (1): Show first that $H_0 = \widehat{G}_0$ and $H_f = \widehat{G}_f$. If $h \in \widehat{G}_0$ then $\iota(h) \approx \mathbf{1}$, and so $\iota(h)(\xi) \approx 1$ for all $\xi \in \text{nst}(*\mathfrak{G})$ by 7.4.9. Consequently, $\iota(h)(j(g)) \approx 1$ for all $g \in G_f$, implying that $h(g) \approx 1$ by 7.4.6 (2). Hence, $h \in H_0$. Thus, $\widehat{G}_0 \subset H_0$ while the reverse inclusion is the claim of 7.4.6 (1).

The inclusion $\widehat{G}_f \subset H_f$ is an easy consequence of 7.2.6 (2) and 7.4.9. The reverse inclusion has a somewhat more complicated proof.

Take $h \in H_f$. Then $\tilde{h} \in G^{\#\wedge}$ (recall that $\tilde{h}(g^\#) = {}^\circ h(g)$). Given ι , determine $\bar{\iota}$ in exactly the same manner as \bar{j} is determined from j . Now, $\bar{\iota} : \widehat{G}^\# \rightarrow \widehat{\mathfrak{G}}$ is a topological isomorphism by 7.4.3. Put $\varkappa := \tilde{h} \circ \bar{j}^{-1} \in \widehat{\mathfrak{G}}$. Since $\widehat{G}^\# = \widehat{G}_f / \widehat{G}_0$, there is some h_1 in \widehat{G}_f satisfying $\bar{\iota}(h_1^\#) = \tilde{h} \circ \bar{j}^{-1}$. If $g \in G_f$ then $\bar{\iota}(h_1^\#)(\bar{j}(g^\#)) = \text{st } \iota(h_1)(\text{st } j(g)) \approx \iota(h_1)j(g) \approx h_1(g)$ by 7.4.6 (2). On the other hand,

$$\bar{\iota}(h_1^\#)(\bar{j}(g^\#)) = \tilde{h}\bar{j}^{-1}(\bar{j}(g^\#)) = \tilde{h}(g^\#) \approx h(g).$$

Thus, $h(g) \approx h_1(g)$ for all $g \in G_f$. This means that $h \cdot \bar{h}_1 \in H_0 = \widehat{G}_0 \subset \widehat{G}_f$. Since $h_1 \in \widehat{G}_f$, it follows that $h \in \widehat{G}_f$, which proves the second of the equalities in question. Item (1) is now clear since (G, j) is a nice hyperapproximant.

(2): This is immediate from (3).

(3): Let $\gamma : L_2(G^\#, \mu) \rightarrow \mathcal{L}_{2,\Delta}(G)^\#$ stand for the embedding induced by the natural homomorphism $\# : G_f \rightarrow G^\#$. In more detail, γ sends each function $f \in L_2(G^\#, \mu_\Delta)$ to the coset of an $\mathcal{L}_{2,\Delta}(G_f)$ -lifting of f , i.e., of $f \circ \#$ in $\mathcal{L}_{2,\Delta}(G)^\#$.

Similarly, let $\tilde{\gamma} : L_2(G^{\#\wedge}, \mu_{\widehat{\Delta}}) \rightarrow \mathcal{L}_{2,\widehat{\Delta}}(\widehat{G})^\#$ stand for the embedding induced by the natural homomorphism $\widehat{\#} : \widehat{G}_f \rightarrow \widehat{G}^\#$. Since (G, j) is a nice hyperapproximant; therefore, the equalities we have just established (1) show that $\widehat{G}^\#$ is canonically isomorphic to $G^{\#\wedge}$, with the isomorphism associating the character \tilde{h} with an element $h^\# \in G^{\#\wedge}$ for $h \in \widehat{G}_f = H_f$. Moreover, recalling the definition of

admissible triple in 7.2.13, we easily infer that the following diagram commutes:

$$\begin{array}{ccc} L_2(G^\#, \mu_\Delta) & \xrightarrow{\mathcal{F}_\Delta^{G^\#}} & L_2(G^{\#\wedge}, \mu_\Delta^\#) \\ \gamma \downarrow & & \downarrow \widehat{\gamma} \\ \mathcal{L}_{2,\Delta}(G)^\# & \xrightarrow{(\Phi_\Delta^G)^\#} & \mathcal{L}_{2,\Delta}(\widehat{G})^\# \end{array}$$

Since the topological isomorphisms \bar{j} and \bar{i} carry the measures μ_Δ on $G^\#$ and $\mu_{\widehat{\Delta}}$ on $G^{\#\wedge}$ to the measures $\bar{\mu}_\Delta$ on \mathfrak{G} and $\bar{\mu}_{\widehat{\Delta}}$ on $\widehat{\mathfrak{G}}$, respectively. Therefore, we arrive at the isomorphisms $j_* : L_2(\mathfrak{G}, \bar{\mu}_\Delta) \rightarrow L_2(G^\#, \mu_\Delta)$ and $i_* : L_2(\widehat{\mathfrak{G}}, \bar{\mu}_{\widehat{\Delta}}) \rightarrow L_2(G^{\#\wedge}, \mu_\Delta^\#)$ such that $j_*(f) := f \circ j$ and $i_*(\varphi) := \varphi \circ i$.

Furthermore, the diagram commutes:

$$\begin{array}{ccc} L_2(\mathfrak{G}, \bar{\mu}_\Delta) & \xrightarrow{\mathcal{F}} & L_2(\widehat{\mathfrak{G}}, \bar{\mu}_{\widehat{\Delta}}) \\ j_* \downarrow & & \downarrow i_* \\ L_2(G^\#, \mu_\Delta) & \xrightarrow{\mathcal{F}_\Delta^{G^\#}} & L_2(G^{\#\wedge}, \mu_\Delta^\#) \end{array}$$

It follows immediately from the definitions that $\gamma \circ j_*(f)$ is the coset of an $\mathcal{S}_{2,\Delta}$ -lifting of f , i.e., $\gamma \circ j_*(f) = j_{2,\Delta}(f)$. Similarly, $\widehat{\gamma} \circ i_* = j_{2,\widehat{\Delta}}$, and $j_{2,\Delta}$ and $j_{2,\widehat{\Delta}}$ are induced by j and i , respectively. Comparing the preceding diagrams, we see that the following diagram commutes

$$\begin{array}{ccc} L_2(\mathfrak{G}, \bar{\mu}_\Delta) & \xrightarrow{\mathcal{F}} & L_2(\widehat{\mathfrak{G}}, \bar{\mu}_{\widehat{\Delta}}) \\ j_{2,\Delta} \downarrow & & \downarrow j_{2,\widehat{\Delta}} \\ \mathcal{L}_{2,\Delta}(G^\#, \mu_\Delta) & \xrightarrow{(\Phi_\Delta^G)^\#} & \mathcal{L}_{2,\Delta}(\widehat{G}^{\#\wedge}) \end{array}$$

yielding (3). \triangleright

7.4.11. We note first that since the couple (G, j) in 7.4.1 is a nonstandard object, Nelson's algorithm is not applicable to the proposition “ (G, j) is a hyperapproximant to \mathfrak{G} ,” because this algorithm deals only with the formulas containing only standard parameters. To obviate this obstacle, we will proceed with the following definition.

A standard sequence $((G_n, j_n))_{n \in \mathbb{N}}$, with G_n a finite abelian group and j_n a mapping of G_n into \mathfrak{G} for $n \in \mathbb{N}$, is a *sequential approximant* or *approximating*

sequence to a separable locally compact abelian group \mathfrak{G} provided that (G_N, j_N) is a hyperapproximant to \mathfrak{G} for all $N \approx +\infty$.

Let $((\widehat{G}_n, \widehat{j}_n))_{n \in \mathbb{N}}$ be a sequential approximant to $\widehat{\mathfrak{G}}$. We call it a *dual* for $((G_n, j_n))_{n \in \mathbb{N}}$ provided that $(\widehat{G}_N, \widehat{j}_N)$ is a dual hyperapproximant to (G_N, j_N) for all $N \approx +\infty$.

A function $f : \mathfrak{G} \rightarrow \mathbb{C}$ is *rapidly decreasing* or has *rapid decay* with respect to a sequential approximant $((G_n, j_n))_{n \in \mathbb{N}}$ to \mathfrak{G} provided that for each relatively compact neighborhood U about the zero of \mathfrak{G} and each infinite $N \in {}^*\mathbb{N}$ the condition holds (cf. 7.4.8)

$$(\forall B \in {}^*\mathcal{P}(G)) \left(B \subset G - G_f \rightarrow \Delta \sum_{g \in G} |{}^*f(j(g))| \approx 0 \right),$$

with $\Delta := |j_N^{-1}({}^*U)| \cdot |G_n|^{-1}$.

The Nelson algorithm fully applies to these definitions. However, it is impossible to obtain each hyperapproximant from some sequential approximant particularly in case the nonstandard universe ${}^*V(\mathbb{R})$ is not an ultrapower of $V(\mathbb{R})$ with respect to some ultrafilter on \mathbb{N} . However, a thorough inspection of the proof of Theorem 7.4.4 shows that there is some sequential approximant to each separable locally compact abelian group.

In the propositions to follow \mathcal{K} and $\widehat{\mathcal{K}}$ stand for the families of all compact subsets of \mathfrak{G} and $\widehat{\mathfrak{G}}$, while \mathbf{T}_0 and $\widehat{\mathbf{T}}_0$ are some bases for the neighborhood filters about the zeros of \mathfrak{G} and $\widehat{\mathfrak{G}}$, respectively.

7.4.12. Assume that to each $n \in \mathbb{N}$ there are a sequence of finite abelian groups G_n and a mapping $j_n : G_n \rightarrow \mathfrak{G}$, with \mathfrak{G} a separable locally compact abelian group and \mathbf{T}_0 a base of relatively compact neighborhood filter about the zero of \mathfrak{G} . Then $((G_n, j_n))_{n \in \mathbb{N}}$ is a sequential approximant to \mathfrak{G} if and only if the following hold:

- (1) $(\forall \xi \in \mathfrak{G})(\forall U \in \mathbf{T}_0)(\exists f \in \prod_{n \in \mathbb{N}} G_n)(\exists n_0 \in \mathbb{N})(\forall n > n_0)$
 $((\xi - j_n(f_n)) \in U);$
- (2) $(\forall K \subset \mathcal{K})(\forall U \in \mathbf{T}_0)(\exists m \in \mathbb{N})(\forall n > m)(\forall g, h \in G_n)$
 $((j_n(g), j_n(h)) \in K \rightarrow (j_n(g + h) - j_n(g) - j_n(h)) \in U)$
 $\wedge (j_n(g) + j_n(-g)) \in U).$

◁ This is immediate by the Nelson algorithm on considering that $\text{nst}({}^*\mathfrak{G}) = \bigcup \{{}^*K : K \subset \mathfrak{G} \text{ is compact}\}$. ▷

7.4.13. Let $((G_n, j_n))_{n \in \mathbb{N}}$ be a sequential approximant to the separable locally compact abelian group \mathfrak{G} . Then the following hold:

- (1) $f : \mathfrak{G} \rightarrow \mathbb{C}$ is a rapidly decreasing function with respect to this sequence if and only if

$$(\forall U \in \mathbf{T}_0)(\forall \varepsilon > 0)(\exists n_0 \in \mathbb{N})(\exists K \in \mathcal{K})(\forall n > n_0) \\ (\forall B \subset J_n^{-1}(\mathfrak{G} - K)) \frac{1}{|J_n^{-1}(U)|} \sum_{g \in B} |f(J_n(g))| < \varepsilon;$$

(2) To each Haar measure μ on \mathfrak{G} there is some U in \mathbf{T}_0 satisfying

$$\int f d\mu = \lim_{n \rightarrow \infty} \frac{1}{|J_n^{-1}(U)|} \sum_{g \in G_n} |f(J_n(g))|$$

for all bounded μ -almost everywhere continuous function $f : \mathfrak{G} \rightarrow \mathbb{C}$ of rapid decay with respect to $((G_n, J_n))_{n \in \mathbb{N}}$.

◁ To prove (1) apply the Nelson algorithm to the definition of rapidly increasing function. To prove (2), do the same with 7.4.8. ▷

7.4.14. Let $((\widehat{G}_n, \widehat{J}_n))_{n \in \mathbb{N}}$ and $((G_n, J_n))_{n \in \mathbb{N}}$ be the same as in 7.4.11. Then $((\widehat{G}_n, \widehat{J}_n))_{n \in \mathbb{N}}$ is dual to $((G_n, J_n))_{n \in \mathbb{N}}$ if and only if the following hold:

- (1) $(\forall V \in \widehat{\mathbf{T}}_0)(\exists n_0 \in \mathbb{N})(\exists K \in \mathcal{K})(\exists \varepsilon > 0)$
 $(\forall n > n_0)(\forall \chi \in \widehat{G}_n)(\forall g \in J_n^{-1}(K)(|\chi(g) - 1| < \varepsilon \rightarrow \widehat{J}_n(\chi) \in V);$
- (2) $(\forall K \in \mathcal{K})(\forall L \in \widehat{\mathcal{K}})(\forall \varepsilon > 0)(\exists n_0 \in \mathbb{N})(\forall n > n_0)$
 $(\forall g \in J_n^{-1}(K))(\forall \chi \in \widehat{J}_n^{-1}(L)(|\widehat{J}_n(\chi)(J_n(g)) - \chi(g)| < \varepsilon).$

◁ To prove, apply the Nelson algorithm to definition of dual hyperapproximant in 7.4.6. ▷

7.4.15. Let $((G_n, J_n))_{n \in \mathbb{N}}$ be a sequential approximant to a separable locally compact abelian group \mathfrak{G} and let $((\widehat{G}_n, \widehat{J}_n))_{n \in \mathbb{N}}$ be a dual sequential approximant to $\widehat{\mathfrak{G}}$. Suppose that μ is a Haar measure on \mathfrak{G} and $\mathcal{F} : L_2(\mathfrak{G}) \rightarrow L_2(\widehat{\mathfrak{G}})$ is the Fourier transform on \mathfrak{G} . Suppose further that $U \in \mathbf{T}_0$ corresponds to μ in accord to 7.4.13 (2).

In this event if f and $|\mathcal{F}(f)|$ are bounded and Haar almost everywhere continuous functions with $|f|^2$ and $|\mathcal{F}(f)|^2$ of rapid decay with respect to the above sequential approximants then

$$\lim_{n \rightarrow \infty} \frac{|J_n^{-1}(U)|}{|G_n|} \sum_{\chi \in \widehat{G}_n} \left| \int_{\mathfrak{G}} f(\xi) \widehat{J}_n(\chi)(\xi) d\mu(\xi) \right. \\ \left. - \frac{1}{|J_n^{-1}(U)|} \sum_{g \in G_n} f(J_n(g)) \chi(g) \right|^2 = 0.$$

◁ To prove, apply the Nelson algorithm to Theorem 7.4.10. ▷

7.4.16. COMMENTS.

(1) The results of this section belong to Gordon [140, 142, 144, 146].

(2) If \mathfrak{G} is a compact group then $G_f = G$ and each standard finite-rank S -continuous unitary representation of G determines (as explained in the previous section) a unitary representation \tilde{T} of $G^\#$ of the same rank. This representation gives rise to an equivalent representation $\tilde{T} \circ \bar{j}$ of \mathfrak{G} . Theorem 7.3.10 shows that if \mathfrak{G} possesses a hyperapproximant (G, j) then each irreducible unitary representation of \mathfrak{G} takes the same shape with some S -continuous irreducible unitary representation T of G .

(3) The main results of this section are stated for separable locally compact abelian groups. However, we may omit the separability assumption in most cases on assuming that the λ^+ -saturation rather than ω^+ -saturation of the nonstandard universe, with λ the *weight* of \mathfrak{G} (which is by definition the least of the cardinals of bases for the topology of \mathfrak{G}).

(4) We may construct a dual approximant to the character group in the cases of the unit circle and a discrete group. This, together with a thorough inspection of the proof of Theorem 7.4.4, shows that each separable locally compact abelian group with a compact and open subgroup possesses a couple of dual hyperapproximants. Therefore, the same holds for all locally compact abelian groups since we may construct dual hyperapproximants to \mathbb{R} straightforwardly (cf. 7.1).

(5) If $|f|^2$ obeys the conditions of 7.4.8, while $|\mathcal{F}(f)|^2$ meets the same conditions on replacing $\mathfrak{G}, G, \Delta, G_f$ with $\widehat{\mathfrak{G}}, \widehat{G}, \widehat{\Delta}, \widehat{G}_f$, respectively; then 7.4.10 (3) amounts to the following:

$$(\Delta|G|)^{-1} \sum_{h \in \widehat{G}} |\mathcal{F}(f)(i(h)) - \Phi_\Delta(*f \circ j)(h)|^2 \approx 0,$$

which reads in more detail as:

$$(\Delta|G|)^{-1} \sum_{h \in \widehat{G}} \left| \int_{*\mathfrak{G}} *f(\xi) \cdot \overline{i(h)(\xi)} d\mu_\Delta(\xi) - \Delta \sum_{g \in G} *f(j(g)) \cdot \overline{h(g)} \right|^2 \approx 0.$$

7.5. Examples of Hyperapproximation

We now consider hyperapproximation of particular groups: the additive group of the reals \mathbb{R} , the unit circle, the τ -adic solenoid, the additive group of τ -adic integers, and profinite abelian groups.

7.5.1. We start with returning the nice hyperapproximant to the additive group of \mathbb{R} as given in 7.4.2. We let $G := \{-L, \dots, L\}$ stand for the additive group of the ring ${}^*\mathbb{Z}/N{}^*\mathbb{Z}$, where $N := 2L + 1$, $N\Delta \approx +\infty$, $\Delta \approx 0$, and $j : G \rightarrow {}^*\mathbb{R}$ acts as $j(k) := k\Delta$. In this case the dual group \widehat{G} is isomorphic to G as shown by assigning to each $n \in G$ the character χ_n with $\chi_n(m) := \exp(2\pi im/N)$. The group $\widehat{\mathbb{R}}$ is isomorphic to \mathbb{R} as shown by assigning to each $t \in \mathbb{R}$ the character \varkappa_t with $\varkappa_t(x) := \exp(2\pi itx)$. A dual hyperapproximant (\widehat{G}, ι) is defined by $\iota(n) := \frac{n}{N\Delta}$, or, more precisely, $\iota(\chi_n)(x) := \exp(\frac{2\pi in}{N\Delta}x)$.

From 7.2.1 (1) we infer 7.4.1 (1). It is easy to verify 7.4.1 (2). Indeed, $j(m) = m\Delta \approx x$ and $\iota(m) = \frac{m}{N\Delta} \approx t$, so that

$$\exp(2\pi itx) = \exp 2\pi i \iota(m)(j(n)) \approx \exp(2\pi imn/N).$$

The corresponding hyperapproximant to the Fourier transform was thoroughly inspected in 7.1.

7.5.2. We now construct a hyperapproximant to the unit circle S , also denoted by S^1 . It is convenient to represent S^1 in the form of the interval $[-1/2, 1/2)$ as was done before. We consider addition modulo 1 as the group operation $+_S$ on S .

The dual group \widehat{S} is isomorphic to the additive group \mathbb{Z} by assigning to $n \in \mathbb{Z}$ the character $\varkappa_n(x) := \exp(2\pi inx)$.

We take as G the same group $\{-L, \dots, L\}$ as in 7.5.1, where $N := 2L+1 \approx +\infty$, and we define the mapping $j : G \rightarrow {}^*S$ by $j(m) := m/N$ for $m \in G$. The mapping $\iota : \widehat{G} \rightarrow {}^*\mathbb{Z}$ is defined on the dual hyperapproximant (\widehat{G}, ι) by $\iota(n) := n$; more precisely, $\iota(\chi_n) := \varkappa_n$, where χ_n is the character defined in 7.5.1. Since \mathbb{Z} is a discrete group, $\widehat{G}_0 = 0$ and $\widehat{G}_f = \mathbb{Z}$. Since S is compact, it suffices to verify only 7.4.1 (2) which is just as trivial as before. Applying Theorem 7.4.10 to this case, we come to the following:

(1) *Let $f : [-1/2, 1/2) \rightarrow \mathbb{C}$ be a Riemann-integrable function. Then*

$$\sum_{n=-L}^L \left| \int_{-1/2}^{1/2} f(x) \exp(-2\pi inx) dx - \frac{1}{N} \sum_{m=-L}^L f(m/N) \exp(-2\pi imn/N) \right|^2 \approx 0$$

whenever $N := 2L + 1$ is an infinite hypernatural.

By a suitable change of variables we come to the following:

(2) *Assume that f is a Riemann-integrable function on $[-l, l]$, while N and Δ satisfy $N := 2L + 1 \approx +\infty$ and ${}^\circ(N\Delta) = 2l$. Then*

$$\sum_{n=-L}^L \left| \int_{-l}^l f(x) \exp(\pi inx/l) dx - \Delta \sum_{m=-L}^L f(m\Delta) \exp(-2\pi imn/N) \right|^2 \approx 0.$$

7.5.3. In the following two subsections we construct hyperapproximants to profinite abelian groups.

So, we consider a standard sequence $((K_n, \varphi_n))_{n \in \mathbb{N}}$, where K_n is a finite abelian group and $\varphi_n : K_{n+1} \rightarrow K_n$ is an epimorphism for all $n \in \mathbb{N}$.

Let $(K, \psi) := \varprojlim (K_n, \varphi_n)$ be the *projective limit* of $((K_n, \varphi_n))_{n \in \mathbb{N}}$. This means that K is a group and there is a sequence of surjective homomorphisms $\psi := (\psi_n)_{n \in \mathbb{N}}$, with $\psi_n : K \rightarrow K_n$ such that $\varphi_n \circ \psi_{n+1} = \psi_n$ for all $n \in \mathbb{N}$. The topology of (K, ψ) is induced from $\prod_n K_n$. Given $N \approx +\infty$, put $G := {}^*K_N$. Then ${}^*\psi_N : K \rightarrow {}^*K_N = G$ is an epimorphism.

(1) Let $j : G \rightarrow {}^*K$ be an internal mapping (not a homomorphism in general) that is a right inverse of ${}^*\psi_N$ with $j(0) = 0$. Then (G, j) is a hyperapproximant to K .

◁ By the definition of the topology of K , we have the following description for infinite proximity on *K :

$$(\forall \alpha, \beta \in {}^*K)(\alpha \overset{K}{\approx} \beta \leftrightarrow (\forall^{\text{st}} n)({}^*\psi_n(\alpha) = {}^*\psi_n(\beta))).$$

The group K is compact, so 7.4.1 (1) is satisfied and we are left with showing that 7.4.1 (2) and 7.4.1 (3), because 7.4.1 (4) holds by definition. Given $n > m$, we introduce the homomorphism $\varphi_{nm} : K_n \rightarrow K_m$ by the rule $\varphi_{nm} := \varphi_{n-1} \circ \dots \circ \varphi_m$. Then $\varphi_{n,n-1} = \varphi_{n-1}$ and $\varphi_{nm} \circ \psi_n = \psi_m$. Hence, given a standard n in \mathbb{N} , we successively infer

$$\begin{aligned} {}^*\psi_n(j(a+b)) &= {}^*\varphi_{Nn} \circ {}^*\psi_N(j(a+b)) \\ &= {}^*\varphi_{Nn}(a+b) = {}^*\varphi_{Nn}(a) + {}^*\varphi_{Nn}(b). \end{aligned}$$

Similarly,

$${}^*\psi_n(j(a) + j(b)) = {}^*\psi_n(j(a)) + {}^*\psi_n(j(b)) = {}^*\varphi_{Nn}(a) + {}^*\varphi_{Nn}(b),$$

which yields 7.4.1 (2) on considering the above description of infinite proximity on *K . By analogy we derive 7.4.1 (3) so completing the proof. ▷

(2) In the context of (1)

$$G_0 = \{a \in G : (\forall^{\text{st}} n)({}^*\varphi_{Nn}(a) = 0)\},$$

with $\varphi_{nm} : K_n \rightarrow K_m$, and $K \simeq G/G_0 = G^\#$.

7.5.4. If $(K, \psi) := \varprojlim (K_n, \varphi_n)$ then the dual \widehat{K} takes the shape $\varprojlim (\widehat{K}_n, \widehat{\varphi}_n)$, with $\widehat{\varphi}_n : \widehat{K}_n \rightarrow \widehat{K}_{n+1}$ acting by the rule $\widehat{\varphi}_n(\chi) := \chi \circ \varphi_n$ ($\chi \in \widehat{K}_n$).

The embeddings $\widehat{\psi}_n : \widehat{K}_n \rightarrow \widehat{K}$ are defined similarly. Given $n > m$, we define the embeddings $\widehat{\varphi}_{nm} : \widehat{K}_m \rightarrow \widehat{K}_n$ as $\widehat{\varphi}_{nm}(\chi) := \chi \circ \varphi_{nm}$ for all $\chi \in \widehat{K}_m$, which yields $\widehat{\varphi}_{nm} = \widehat{\varphi}_m \circ \cdots \circ \widehat{\varphi}_{n-1}$ and $\widehat{\psi}_m = \widehat{\psi}_n \circ \widehat{\varphi}_{nm}$.

(1) Assume that (G, j) is the hyperapproximant of K defined as in 7.5.3(1); i.e., $G := K_N$ and $j : G \rightarrow {}^*K$ is the right inverse of ${}^*\psi_N$. Then $(\widehat{K}, \widehat{\psi}_N)$ is a hyperapproximant of \widehat{K} dual to (G, j) .

◁ Note first that 7.4.1(2–4) hold automatically because $\widehat{\psi}_N$ is a homomorphism. Since \widehat{K} is the inductive limit of (\widehat{K}_n) ; therefore, $\widehat{K} = \bigcup_{n \in \mathbb{N}} A_n$, with $A_n := \{\chi \circ \psi_n : \chi \in \widehat{K}_n\}$. By transfer, ${}^*\widehat{K} = \bigcup_{n \in {}^*\mathbb{N}} A_n$. Since \widehat{K}_n is a standard finite set, it follows that ${}^*A_n = \{\chi \circ {}^*\psi_n : \chi \in \widehat{K}_n\}$. Thus, each standard element $\varkappa \in {}^*\widehat{K}$ has the shape $\varkappa = \chi \circ {}^*\psi_n$ for some standard n and $\chi \in \widehat{K}_n$. Consequently, ${}^*\widehat{\varphi}_{nM}(\chi) \in \widehat{K}_n$ and $\widehat{\psi}_N({}^*\widehat{\varphi}_{nN}(\chi)) = \varkappa$, which yields 7.4.1(1). We have 7.4.6(1) automatically because K is compact. If $\chi \in \widehat{K}_n$, then ${}^*\widehat{\psi}_N(\chi)(j(a)) = \chi({}^*\psi_N(j(a))) = \chi(a)$, yielding 7.4.6(2). ▷

(2) In the context of 7.5.3(1) and (1), assume that $f : K \rightarrow \mathbb{C}$ is a bounded and Haar almost everywhere continuous function. Then

$$\sum_{\chi \in \widehat{K}_N} \left| \int_{{}^*K} f(\alpha) \overline{\chi(\psi_N(\alpha))} d^* \mu_K(\alpha) - |K_N|^{-1} \sum_{a \in K_N} f(j(a)) \overline{\chi(a)} \right|^2 \approx 0,$$

with μ_K standing for the Haar measure on K satisfying $\mu_K(K) = 1$.

◁ Straightforward from 7.4.10 and 7.4.16(5). ▷

7.5.5. We now apply the results of the preceding subsection to the constructing a hyperapproximant to the ring Δ_τ of τ -adic integers (see [151, 174]).

We let the symbol $a \mid b$ denote the fact that b divides a , while $\text{rem}(a, b)$ will stand for the remainder from dividing a by b .

Let $\tau := (a_n)_{n \in \mathbb{N}}$ be a standard sequence of naturals such that $a_n > 1$ and $a_n \mid a_{n+1}$. Denote by A_n the ring $\mathbb{Z}/a_n\mathbb{Z}$ which in our case is conveniently regarded as the ring of smallest positive residues modulo a_n , i.e., $A_n = \{0, 1, \dots, a_n - 1\}$.

Let $\varphi_n : A_{n+1} \rightarrow A_n$ be the epimorphism, sending $a \in A_{n+1}$ to the remainder from dividing a by a_n ; in symbols, $\varphi_n(a) := \text{rem}(a, a_n)$. The ring $\Delta_\tau := \varprojlim (A_{n+1}, \varphi_n)$ is the *ring of τ -adic integers*.

We define the embedding $\nu : \mathbb{Z} \rightarrow \Delta_\tau$ by putting $\nu(a)_n := \text{rem}(a, a_n)$ for all $a \in \mathbb{N}$ and $n \in \mathbb{N}$.

Then the sequence $\nu(a)$ belongs to $\prod_{n \in \mathbb{N}} A_n$ and $\varphi_n(\nu(a)_{n+1}) = \nu(a)_n$, and so $\nu(a) \in \Delta_\tau$. It is easy that $\nu(\mathbb{Z})$ is dense in Δ_τ . Put $j_n := \nu|_{A_n} : A_n \rightarrow \Delta_\tau$. Then j_n is a right inverse of $\psi_n : \Delta_\tau \rightarrow A_n$. Indeed, if $\xi := (\xi_n)_{n \in \mathbb{N}} \in \Delta_\tau$ then $\psi_n(\xi) = \xi$. Since $\text{rem}(a, a_n) = a$ for $a \in A_n$, infer that $\psi_n(\nu(a)) = a$.

If $N \approx +\infty$ then the couple $({}^*A_N, {}^*j_N)$ is a hyperapproximant to Δ_τ . Moreover, Δ_τ is topologically isomorphic to ${}^*A_N/G_0$, with

$$G_0 := \{a \in {}^*A_N : (\forall^{\text{st}} n)(a_n | a)\}.$$

◁ Immediate from 7.5.3 (1, 2). ▷

7.5.6. We now describe the dual $\widehat{\Delta}_\tau$ (see [174]). Assign $\mathbb{Q}^{(\tau)} := \{\frac{m}{a_n} : m \in \mathbb{Z}, n \in \mathbb{N}\}$. Since $a_n | a_m$ for $n \leq m$, it follows that $\mathbb{Q}^{(\tau)}$ is a subgroup of the additive group \mathbb{Q} . Obviously, $\mathbb{Z} \subset \mathbb{Q}^{(\tau)}$. Put $\mathbb{Z}^{(\tau)} := \mathbb{Q}^{(\tau)}/\mathbb{Z}$.

It is well known that $\widehat{\Delta}_\tau \cong \mathbb{Z}^{(\tau)}$. To describe an isomorphism we introduce the following notation. If $\xi := (\xi_n)_{n \in \mathbb{N}} \in \Delta_\tau$ then we write $\xi_n = \text{rem}(\xi, a_n)$. This agrees with the case in which $\xi = \nu(a)$ for some $a \in \mathbb{Z}$; i.e., $\text{rem}(a, a_n) = \text{rem}(\nu(a), a_n)$.

In what follows, we identify $\nu(a)$ with a , implying that $\mathbb{Z} \subset \Delta_\tau$. Then the equality $\xi = \eta a_n + \text{rem}(\xi, a_n)$ holds for some $\eta \in \Delta_\tau$.

Put $\{\xi/a_n\} := (\text{rem}(\xi, a_n))/a_n \in \mathbb{Q}^{(\tau)}$. It is then easy that $\{C\xi/a_n\} \equiv C\{\xi/a_n\} \pmod{\mathbb{Z}}$ and $\{\xi/a_n + \eta\} = \{\xi/a_n\}$, with $C \in \mathbb{Z}$ and $\eta \in \Delta_\tau$.

Assume now that (C/a_n) is the coset of $C/a_n \in \mathbb{Q}^{(\tau)}$ in $\mathbb{Z}^{(\tau)}$. Then the character $\chi_{(C/a_n)} \in \widehat{\Delta}_\tau$ is given by the formula

$$\chi_{(C/a_n)}(\xi) = \exp(2\pi i \{\xi/a_n\}) \quad (\xi \in \Delta_\tau).$$

We now describe the embedding $\widehat{\psi}_n : \widehat{A}_n \rightarrow \widehat{\Delta}_\tau$.

In the case under consideration \widehat{A}_n is isomorphic to A_n by sending each m to the character $\chi_m \in \widehat{A}_n$ given by $\chi_m(a) := \exp(2\pi i m a/a_n)$ for $a \in A_n$.

Hence,

$$\begin{aligned} \widehat{\psi}_n(\chi_m)(\xi) &= \chi_m(\psi_n(\xi)) = \chi_n(\text{rem}(\xi, a_n)) \\ &= \exp(2\pi i m \text{rem}(\xi, a_n)/a_n) = \exp(2\pi i \{m\xi/a_n\}). \end{aligned}$$

Carrying out appropriate identifications, we may assume that $\widehat{\psi}_n : A_n \rightarrow \mathbb{Z}^{(\tau)}$ is given by the formula $\widehat{\psi}_n(m) = (m/a_n)$.

For an arbitrary $N \approx +\infty$, the couple $({}^*A_N, {}^*\widehat{\psi}_N)$ is a hyperapproximant of $\widehat{\Delta}_\tau$ dual to $({}^*A_N, {}^*j_N)$.

◁ Follows from 7.4.10. ▷

(1) If $N \approx +\infty$ and $f : \Delta_\tau \rightarrow \mathbb{C}$ is a standard bounded function that is continuous almost everywhere with respect to the Haar measure μ_τ , with $\mu_\tau(\Delta_\tau) = 1$, then

$$\sum_{k=0}^{*a_N-1} \left| \int_{*\Delta_\tau} *f(\xi) \exp(-2\pi i \{k\xi/*a_N\}) d\mu_\tau(\xi) - a_N^{-1} \sum_{n=0}^{*a_N-1} *f(n) \exp(-2\pi i kn/*a_N) \right|^2 \approx 0.$$

◁ Follows from 7.4.4 (5). ▷

7.5.7. In the forthcoming two subsections we hyperapproximate the τ -adic solenoid. Recall (see [151, 174]) that the τ -adic solenoid Σ_τ is represented as $[0, 1) \times \Delta_\tau$, with the group operation $+_\tau$ defined as

$$(x, \xi) +_\tau (y, \eta) := (\{x + y\}, \xi + \eta + [x + y]),$$

with $[a]$ and $\{a\}$ standing for the integral and fractional parts of a real a . The topology on Σ_τ is given by the system $(V_n)_{n \in \mathbb{N}}$ of neighborhoods about zero, with

$$V_n := \{(x, \xi) : 0 \leq x < 1/a_n, (\forall k \leq n)(\text{rem}(x, a_k) = 0)\} \cup \{(x, \xi) : 1 - 1/a_n < x < 1, (\forall k \leq n)(\text{rem}(x + 1, a_k) = 0)\}$$

(recall that $\tau := \{a_n : n \in \mathbb{N}\}$).

The above readily yields the following description for the infinitesimals in Σ_τ .

(1) If $(x, \xi) \in * \Sigma_\tau$ then

$$(x, \xi) \stackrel{\Sigma_\tau}{\approx} 0 \leftrightarrow x \approx 0 \wedge \xi \stackrel{\Delta_\tau}{\approx} 0 \vee x \approx 1 \wedge \xi + 1 \stackrel{\Delta_\tau}{\approx} 0.$$

Distinguishing some $N \approx +\infty$, put $G := * \mathbb{Z} / * a_N^2 * \mathbb{Z} = \{0, 1, \dots, * a_N^2 - 1\}$ and introduce the mapping $j : G \rightarrow \Sigma_\tau$ by the rule

$$j(a) := (\{a/a_N\}, [a/a_N]) \quad (a \in G).$$

As in the preceding subsection, we assume that $\mathbb{Z} \subset \Delta_\tau$. In this event, $[a/a_N] < a_N$.

(2) The couple (G, j) is a hyperapproximant to the τ -adic solenoid Σ_τ .

◁ Since $a +_G b \equiv a + b \pmod{a_N^2}$; therefore, $a +_G b \equiv a + b \pmod{a_N}$, and so

$$\begin{aligned} \{(a +_G b)/a_N\} &= (\text{rem}(a +_G b, a_N))/a_N \\ &= (\text{rem}(a + b, a_N))/a_N = \{\{a/a_N\} + \{b/a_N\}\}. \end{aligned}$$

We will show that $[(a +_G b)/a_N] \stackrel{\Delta_\tau}{\approx} [(a + b)/a_N] = [a/a_N] + [b/a_N] + [\{a/a_N\} + \{b/a_N\}]$. By (1) this implies that $j(a +_G b) \stackrel{\Sigma_\tau}{\approx} j(a) + j(b)$.

Put $a := q(a)a_N + r(a)$ and $b := q(b)a_N + r(b)$, i.e., $q(a) = [a/a_N]$ and $q(b) = [b/a_N]$. Let $a + b := q(a + b)a_N + r(a + b)$. If $q(a + b) = sa_N + r$ then $a + b = sa_N^2 + ra_N + r(a + b)$ and $ra_N + r(a + b) \leq (a_N - 1)a_N + a_N - 1 = a_N^2 - 1$. Thus, $a +_G b = \text{rem}((a + b), a_N^2) = ra_N + r(a + b)$, and so $q(a +_G b) = r$. Therefore, $q(a + b) \equiv q(a +_G b) \pmod{a_N}$. Since $a_n \mid a_N$ for all standard n , it follows that $q(a + b) \approx q(a +_G b)$.

To prove the relation $j(-_G a) \approx j(a)$ it suffices to represent a as $a = qa_N + r$, use the fact that $-_G a = a_N^2 - a$, and settle the two cases $r = 0$ and $r \neq 0$.

To show that (G, j) is a hyperapproximant to Σ_τ it remains to prove (cf. 7.4.1) that to each couple $(x, \xi) \in \Sigma_\tau$ there is some a in G satisfying $j(a) \approx (x, \xi)$. To this end, choose some $r < a_N$ such that $r/a_N \leq x < (r + 1)/a_N$ and put $q := \text{rem}(*x, a_N)$. Then $q < a_N$, and we are done on letting $a := qa_N + r$. ▷

By 7.5.5 and (1), $\Sigma_\tau \simeq G/G_0$, with

$$G_0 = \{a \in G : \{a/a_N\} \approx 0 \wedge [a/a_N] \stackrel{\Delta_\tau}{\approx} 0 \vee \{a/a_N\} \approx 1 \wedge [a/a_N] + 1 \stackrel{\Delta_\tau}{\approx} 0\}.$$

We may also provide a more concise description for G_0 .

$$(3) \quad G_0 = \{a \in G : (\forall^{\text{st}n})(\exp(2\pi ia/(a_N a_n)) \approx 1)\}.$$

◁ Given a standard n , assume that $\exp(2\pi ia/(a_N a_n)) \approx 1$. Then there is some $k \in \mathbb{Z}$ satisfying $a/(a_N a_n) \approx k$. If $a := qa_N + r$ then $a/(a_N a_n) = (q + r/a_N)/a_n \approx k \in {}^*\mathbb{Z}$. Since a_n is standard, it follows that $q + r/a_N \approx ka_n \in {}^*\mathbb{Z}$, where $q \in {}^*\mathbb{Z}$ and $0 \leq r/a_N < 1$, so that the two cases $r/a_N \approx 0$ and $r/a_N \approx 1$ are possible. In the first case $q \approx ka_n$ and $q = ka_n$, because both numbers are integers. Therefore, $r/a_N \approx 0$ and $q \stackrel{\Delta_\tau}{\approx} 0$, and so $a \in G_0$. In the second case of $r/a_N \approx 1$ if $q \not\equiv -1 \pmod{a_n}$ then $q + 1 = ta_n + s$, where $0 < s < a_n$. Therefore, $q + r/a_N = s - 1 + ta_n + r/a_N \approx s + ta_n \approx ka_n$. Then $s/a_n + t \approx k$, which is impossible if $s \neq 0$. Thus, $q \equiv -1 \pmod{a_n}$, and again $a \in G_0$.

Suppose conversely that $a \in G_0$ and $a := qa_N + r$. The two cases are in order:

(a) $r/a_N \approx 0$, $q \stackrel{\Delta_\tau}{\approx} 0$; (b) $r/a_N \approx 1$, $q \stackrel{\Delta_\tau}{\approx} 0$. In the first case $\exp(2\pi ia/(a_N a_n)) = \exp(2\pi i(a/a_n + r/(a_N a_n))) = \exp(2\pi ia/(a_N a_n)) \approx 1$, since $q/a_n \in {}^*\mathbb{Z}$. The second case is settled similarly. ▷

(4) If τ is a standard sequence $(a_n)_{n \in \mathbb{N}}$ of naturals with $a_n > 1$ and $a_n | a_{n+1}$, $G := \{0, 1, \dots, a_N^2 - 1\}$ is the additive group of ${}^*\mathbb{Z}/a_N^2{}^*\mathbb{Z}$, and $G_0 = \{a \in G : (\forall^{\text{st}} n)(\exp(2\pi i a / (a_N a_n))) \approx 1\}$; then the τ -adic solenoid Σ_τ is topologically isomorphic to $G^\# := G/G_0$.

7.5.8. We now construct a dual hyperapproximant of $\widehat{\Sigma}_\tau$. It is well known that $\widehat{\Sigma}_\tau \simeq \mathbb{Q}^{(\tau)} := \{m/a_n : m \in \mathbb{Z}, n \in \mathbb{N}\}$. We may exhibit an isomorphism by associating with each $\alpha := m/a_n \in \mathbb{Q}^{(\tau)}$ the character $\varkappa_\alpha \in \widehat{\Sigma}_\tau$ such that

$$\varkappa_\alpha(x, \xi) := \exp(2\pi i \alpha(x + \text{rem}(\xi, a_n))) \quad (x \in [0, 1], \xi \in \Delta_\tau)$$

(see [151, 174]).

As in the case of an arbitrary finite group, \widehat{G} is isomorphic to G by sending each $b \in G$ to the character χ_b with $\chi_b(a) := \exp(2\pi i ab/a_N^2)$ for all $a \in G$. As a dual hyperapproximant to the hyperapproximant of Σ_τ constructed above, consider the couple (G, ι) with $\iota : G \rightarrow {}^*\mathbb{Q}^{(\tau)}$ acting by the rule $\iota(b) := b/a_N$. More precisely, $\iota : \widehat{G} \rightarrow \widehat{\Sigma}_\tau$ carries each $\chi_b \in \widehat{G}$ to the character $\varkappa_{b/a_N} \in {}^*\widehat{\Sigma}_\tau$, on assuming that G is presented as the group of absolutely smallest residues, i.e., $G = \{-\frac{1}{2}a_N^2, \dots, \frac{1}{2}a_N^2 - 1\}$.

To check 7.4.1 (1–4) is an easy matter in this case. As regards 7.4.6 (1, 2) it suffices again to validate the second condition, which holds now as equality. Indeed, if $a \in G$ and $a := qa_N + r$ then $[a/a_N] = q < a_N$, i.e., $\text{rem}([a/a_N], a_N) = [a/a_N] = q$. Now $\varkappa_{\iota(b)}(j(a)) = \exp(2\pi i (b/a_N)(r/a_N + q)) = \exp(2\pi i ab/a_N^2) = \chi_b(a)$.

Let $f : \Sigma_\tau \rightarrow \mathbb{C}$ be a bounded almost everywhere continuous function with respect to the Haar measure μ_{Σ_τ} , with $\mu_{\Sigma_\tau}(\Sigma_\tau) = 1$. If $N \approx +\infty$ then

$$(1) \int_{\Sigma_\tau} f d\mu_\tau = \circ \left(\frac{1}{a_N^2} \sum_{k=0}^{a_N^2-1} f(\{k/a_N\}, [k/a_N]) \right);$$

$$(2) \sum_{m=-a_N^2/2}^{a_N^2/2-1} \left| \int_{\Sigma_\tau} f(x, \xi) \exp(-2\pi i (m/a_N)(x + \text{rem}(\xi, a_N))) d\mu_{\Sigma_\tau} - \frac{1}{a_N^2} \sum_{k=0}^{a_N^2-1} f(\{k/a_N\}, [k/a_N]) \exp(-2\pi i km/a_N^2) \right|^2 \approx 0.$$

◁ The claim follows from 7.4.10 and 7.4.16 (5). ▷

7.5.9. We now construct a hyperapproximant to the additive group of the field \mathbb{Q}_p of p -adic numbers, with p a standard prime integer.

Assume given M and N with $M, N, N - M \approx +\infty$. As a hyperfinite abelian group G we consider the additive group of the ring ${}^*\mathbb{Z}/p^N{}^*\mathbb{Z}$, which is conveniently represented here as the group of smallest positive residues $G := \{0, 1, \dots, p^N - 1\}$. We also define the mapping $j : G \rightarrow {}^*\mathbb{Q}_p$, by putting $j(n) := n/p^M \in {}^*\mathbb{Q} \subset \mathbb{Q}_p$ for $n \in G$.

(1) If $n \in G$ then

$$j(n) \in \text{nst}({}^*\mathbb{Q}_p) \leftrightarrow (\exists^{\text{st}} k \in \mathbb{N})(p^{M-k} | n).$$

Moreover, if

$$n = a_{-k}p^{M-k} + a_{-k+1}p^{M-k+1} + \dots + a_{N-M-1}p^{N-1},$$

with $0 \leq a_i < p$, then

$$\tilde{j}(n) = \text{st}(j(n)) = \sum_{i=-k}^{\infty} a_i p^i.$$

(Recall that $\tilde{j} := \text{st} \circ j|_{G_f}$, cf. 7.4.1.)

◁ If $p^{M-k} | n$ then n has the shape of the statement and the claim follows from the fact that $N - M - 1$ is an infinite hypernatural. Conversely, suppose that $np^{-M} \approx \xi \in \mathbb{Q}_p$ and $|\xi|_p = p^k$, with $k \in \mathbb{Z}$. Since $|np^{-M} - \xi|_p \approx 0$, there is an infinite $b \in {}^*\mathbb{N}$ such that $np^{-M} - \xi = p^b \varepsilon_1$ where ε_1 is the unit of ${}^*\mathbb{Z}_p$. By assumption, $\xi = p^{-k} \varepsilon_2$, where ε_2 is the unit of \mathbb{Z}_p . This implies that $n = p^{M-k} \varepsilon_2 + p^{M+b} \varepsilon_1$. Since k is standard; therefore, $M - k < M + b$, implying that $p^{M-k} | n$ in ${}^*\mathbb{Z}_p$ and, hence, also in \mathbb{Z} . ▷

(2) The couple (G, j) is a hyperapproximant to the additive group \mathbb{Q}_p^+ of \mathbb{Q}_p . Moreover,

$$\begin{aligned} G_f &= \{n \in G : (\exists^{\text{st}} k)(p^{M-k} | n)\}, \\ G_0 &= \{n \in G : (\forall^{\text{st}} k)(p^{M+k} | n)\}. \end{aligned}$$

◁ The sought equalities follow from (1). To verify 7.4.1 (1–4) it suffices to show that $\tilde{j} : G_f \rightarrow \mathbb{Q}_p$ is an epimorphism.

Denote the addition of G by \oplus . Let $n = n_1 \oplus n_2$, i.e., $n_1 + n_2 = n + tp^M$. Then $|n_1 p^{-M} + n_2 p^{-M} - np^{-M}|_p \leq p^{N-M}$, and so $\tilde{j}(n_1 + n_2) = \tilde{j}(n_1 \oplus n_2)$ because $N - M$ is infinite.

Since $\text{st} : \text{nst}({}^*\mathbb{Q}_p) \rightarrow \mathbb{Q}_p$ is a homomorphism, obtain $\tilde{j}(n_1 + n_2) = \tilde{j}(n_1) + \tilde{j}(n_2)$. This proves that \tilde{j} is a homomorphism. To verify that \tilde{j} is surjective, note that if $\xi = \sum_{i=-k}^{\infty} a_i p^i$ then, defining n as in (1), we find that $j(n) \approx \xi$. ▷

(3) $\mathbb{Q}_p^+ \simeq G_f/G_0 = G^\#$.

Let $G^{(0)} := \{n \in G : p^M | n\}$. Then $G^{(0)}$ is an internal subgroup of G satisfying $G_0 \subset G^{(0)} \subset G_f$. Since $|G^{(0)}| = p^{N-M}$, we may take $\Delta := p^{M-N}$ as a normalizing factor of the triple (G, G_0, G_f) .

(4) The equality $\tilde{j}(G^{(0)}) = \mathbb{Z}_p$ holds. Moreover, the normalizing factor $\Delta := p^{M-N}$ induces on \mathbb{Q}_p the Haar measure $\bar{\mu}_\Delta$ with $\bar{\mu}_\Delta(\mathbb{Z}_p) = 1$ (this measure is denoted by μ_p in the sequel).

7.5.10. We now derive a standard version of the condition for an $\mathcal{S}_{1,\Delta}$ -integrable lifting to exist as stated in 7.4.8.

(1) For $f : \mathbb{Q}_p \rightarrow \mathbb{C}$ to satisfy the condition (cf. 7.4.8)

$$(\forall B \in {}^* \mathcal{P}(G)) \left(B \subset G - G_f \rightarrow \Delta \sum_{g \in G} |{}^* f(j(g))| \approx 0 \right),$$

it is necessary and sufficient that the relation

$$\lim_{m,n \rightarrow \infty} \frac{1}{p^n} \sum_{\substack{0 \leq k < p^{m+n+l} \\ p^l \nmid k}} |f(k/p^{m+l})| = 0$$

holds uniformly in l .

◁ If $B \subset G - G_f$ then $p^{M-k} \nmid L$ for all $n \in B$ and every standard k . The latter amounts to the condition that $p^{M-L} \nmid n$ for some infinite L . Thus, $B \subset G - B_f$ if and only if $B \subset B_L := \{n : p^{M-L} \nmid n\}$. However, B_L is an internal set for every infinite L . Hence, the hypothesis may be reformulated as follows: If L is an infinite hypernatural such that $L < M < N$ and $N - M$ is also infinite then

$$\frac{1}{p^{N-M}} \sum_{\substack{0 \leq k < p^N \\ p^{M-L} \nmid k}} |{}^* f(k/p^M)| \approx 0.$$

Putting $n := N - M$, $m := L$, and $l := M - L$, infer that

$$\frac{1}{p^n} \sum_{\substack{0 \leq k < p^{m+n+l} \\ p^l \nmid k}} |{}^* f(k/p^{m+l})| \approx 0$$

for all infinite m and n and for every l . The last relation yields the claim. ▷

(2) If $f : \mathbb{Q}_p \rightarrow \mathbb{C}$ is a bounded and Haar almost everywhere continuous function satisfying (1), then f is integrable and

$$\int_{\mathbb{Q}_p} f d\mu_p = \overset{\circ}{\left(\frac{1}{p^{N-M}} \sum_{0 \leq k < p^N} {}^* f(k/p^M) \right)},$$

or, standardly,

$$\int_{\mathbb{Q}_p} f d\mu_p = \lim_{m,n \rightarrow \infty} \frac{1}{p^n} \sum_{0 \leq k < p^{m+n}} {}^* f(k/p^m).$$

7.5.11. We now construct a dual hyperapproximant to $\widehat{\mathbb{Q}}_p^+$. Recall that $\widehat{\mathbb{Q}}_p^+$ is isomorphic with \mathbb{Q}_p^+ by sending with each $\xi \in \mathbb{Q}_p^+$ to the character $\varkappa_\xi(\eta) := \exp(2\pi i\{\xi\eta\})$, with $\{\zeta\}$ the fractional part of a p -adic number ζ . Identifying $\widehat{\mathbb{Q}}_p^+$ and \mathbb{Q}_p^+ in this manner, we construct a hyperapproximant (G, \widehat{j}) to \mathbb{Q}_p by letting $\widehat{j}(n) := n/p^{N-M}$ for $n \in G$.

(1) The couple (G, \widehat{j}) is a hyperapproximant to the dual $\widehat{\mathbb{Q}}_p^+$ of (G, j) .

◁ The claim about hyperapproximation for (G, \widehat{j}) was established in 7.5.9 (2). We are left with checking that (G, \widehat{j}) is dual to (G, j) (see 7.4.6). To verify 7.4.6 (2) is an easy matter:

$$\begin{aligned} \varkappa_{j(m)}(\widehat{j}(n)) &= \exp(2\pi i\{\widehat{j}(n)j(m)\}) = \exp(2\pi i\{nm/p^N\}) \\ &= \exp(2\pi inm/p^N) = \chi_m(n). \end{aligned}$$

To verify 7.4.6 (1), it suffices to show that

$$\begin{aligned} (\forall m)((\exists^{\text{st}}k)(p^{M-k} \mid m) \rightarrow \exp(2\pi imn/p^N) \approx 1) \\ \rightarrow (\forall^{\text{st}}k)(p^{N-M+k} \mid n). \end{aligned}$$

Were this false, we would find some k satisfying $n = qp^{N-M+k} + r$ and $0 < r < p^{N-M+k}$. The two cases are possible.

(1): ${}^\circ(r/p^{N-M+k}) = 0$. Put $a := \lfloor p^{N-M+k}/(2r) \rfloor$ and $m := ap^{M-k}$ (obviously, $m < p^N$). Then

$$\begin{aligned} &\exp(2\pi imn/p^N) \\ &= \exp(2\pi i[p^{N-M+k}/(2r)](r/p^{N-M+k})) \approx \exp(\pi i) = -1, \end{aligned}$$

which is a contradiction.

(2): ${}^\circ(r/p^{N-M+k}) = \alpha$ and $0 < \alpha \leq 1$. Putting $m := p^{M-k-1}$, infer that $\exp(2\pi imn/p^N) \approx \exp(2\pi i\alpha/p) \neq 1$, and we again arrive at a contradiction. ▷

(2) Consider the Fourier transform $\mathcal{F} : L_2(\mathbb{Q}_p) \rightarrow L_2(\mathbb{Q}_p)$ with

$$\mathcal{F}(f)(\xi) := \int_{\mathbb{Q}_p} f(\eta) \exp(-2\pi i\{\xi\eta\}) d\mu_p(\eta).$$

Let $f \in L_2(\mathbb{Q}_p)$ be such that $|f|^2$ and $|\mathcal{F}(f)|^2$ are bounded almost everywhere continuous functions satisfying 7.5.10 (1). Then

$$\begin{aligned} &\frac{1}{p^M} \sum_{k=0}^{p^N-1} \left| \int_{\mathbb{Q}_p} f(\eta) \exp(-2\pi i\{\eta k/p^{N-M}\}) d\mu_p(\eta) \right. \\ &\left. - \frac{1}{p^{N-M}} \sum_{n=0}^{p^N-1} f(n/p^M) \exp(-2\pi ikn/p^N) \right|^2 \approx 0, \end{aligned}$$

whenever $N, M, N - M \approx +\infty$.

◁ The claim follows from Theorem 7.4.10 and 7.4.16 (5). ▷

7.5.12. The field of p -adic numbers is abstracted to the ring \mathbb{Q}_α of α -adic numbers, where $\alpha = \{a_n : n \in \mathbb{Z}\}$ is a doubly infinite sequence of naturals such that $a_n | a_{n+1}$ for $n \geq 0$ and $a_{n+1} | a_n$ for $n < 0$. This ring is described in detail in [151; Chapter 2, §3.7]. It is shown in [174] that the group $\widehat{\mathbb{Q}}_\alpha^+$ is isomorphic to \mathbb{Q}_α^+ , with $\widehat{\alpha}(n) := \alpha(-n)$.

Let $M, N \approx +\infty$, and let $G := \{0, 1, \dots, {}^*a_M {}^*a_N - 1\}$ be the additive group of ${}^*\mathbb{Z}/({}^*a_M {}^*a_N){}^*\mathbb{Z}$. Suppose that $j : G \rightarrow \mathbb{Q}_\alpha^+$ and $\widehat{j} : G \rightarrow \mathbb{Q}_\alpha^+$ are defined by the rules $j(a) := a a_{-M}^{-1}$ and $\widehat{j}(a) := a a_N^{-1}$ for $a \in G$.

Then the couple (G, j) is a hyperapproximant to \mathbb{Q}_α^+ , and the couple $(\widehat{G}, \widehat{j})$ is a dual hyperapproximant to \mathbb{Q}_α^+ . Moreover,

$$\begin{aligned} G_f &:= \{a \in G : (\exists {}^{\text{st}}k \in \mathbb{Z})(a_{-M} a_k^{\text{sgn } k} | a)\}, \\ G_0 &:= \{a \in G : (\forall {}^{\text{st}}k \in \mathbb{Z})(a_{-M} a_k^{\text{sgn } k} | a)\}, \end{aligned}$$

and $\Delta := a_N^{-1}$ is a normalizing factor for the triple (G, G_0, G_f) . Similar definitions apply to $\widehat{G}_f, \widehat{G}_0$, and $\widehat{\Delta} := a_{-M}^{-1}$, a normalizing factor for the triple $(\widehat{G}, \widehat{G}_0, \widehat{G}_f)$.

◁ The proof of this proposition is analogous to 7.5.9–7.5.11. ▷

7.5.13. COMMENTS.

(1) Hyperapproximation of the unit circle (cf. 7.5.2) was studied in detail by Luxemburg in [329]; however, the concept of hyperapproximant was not suggested. The article [329] in particular contains 7.5.2 (2) for continuous functions as well as many interesting applications of infinitesimal analysis to Fourier series. However, the approach of [329] was limited since the theory of Loeb measure and the technique of saturation were not available at that time.

(2) Propositions 7.5.3 (1, 2) remain valid in case K_n is a ring. In this event K and G are rings too, and the mapping j from G to *K is an “almost homomorphism,” i.e., in addition to the already-mentioned properties, j maintains the relation $j(ab) \approx j(a)j(b)$ for all a and b . Moreover, the set G_0 , as defined in 7.5.3 (2), will be a two-sided ideal.

(3) Proposition 7.4.8 shows that, given f satisfying 7.5.6 (1), we have

$$\int_{\Delta_\tau} f(\xi) d\mu_\tau(\xi) = \circ \left(\frac{1}{a_N} \sum_{n=0} {}^*a_N^{-1} {}^*f(n) \right),$$

which amounts to the standard equality

$$\int_{\Delta_\tau} f(\xi) d\mu_\tau(\xi) = \lim_{N \rightarrow \infty} \frac{1}{a_N} \sum_{n=0}^{a_N-1} f(n).$$

(4) If f is continuous then we obtain a stronger equality:

$$\int_{\Delta_\tau} f(\xi) d\mu_\tau(\xi) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} f(n).$$

This follows from the strict ergodicity of the shift by 1 in Δ_τ which ensues in turn from the inequality $\chi(1) \neq 1$ holding for every nontrivial character $\chi \in \widehat{\Delta}_\tau$ (see [245; Chapter 4, § 1, Theorem 1]). In regard to the sequence $\tau := \{(n + 1)! : n \in \mathbb{N}\}$, this equality was derived within analytic number theory for a slightly broader function class containing all functions in 7.5.6 (1) (for instance, see [403]). This result shows that for f to be bounded and almost everywhere continuous the condition is not necessary that $*f \circ j$ be a lifting of f (here f is defined on a compact abelian group \mathfrak{G} , and (G, j) is a hyperapproximant to \mathfrak{G}).

(5) If we take as τ the sequence $(p^{n+1})_{n \in \mathbb{N}}$ then Δ_τ is the ring \mathbb{Z}_p of p -adic integers. Consequently (cf. 7.5.5), $\mathbb{Z}_p \simeq K_p/K_{p0}$, where $K_p = *\mathbb{Z}/p^N*\mathbb{Z}$, $K_{p0} := \{a \in K_p : (\forall^{\text{st}} n)(p^n | a)\}$, and $N \approx +\infty$.

(6) The ring $\Delta := \Delta_{\{(n+1)! : n \in \mathbb{N}\}}$, referred sometimes to as the ring of *polyadic integers*, plays a rather important role in number theory. From 7.5.5 it follows that $\Delta \simeq K/K_0$, with

$$K := *\mathbb{Z}/N!*\mathbb{Z},$$

$$K_0 := \{a \in K : (\forall^{\text{st}} n)(n | a)\}$$

and $N \approx +\infty$.

Hyperapproximation enables us to give a simple proof to the fact that $\Delta \simeq \prod_{p \in P} \mathbb{Z}_p$, with P the set of all prime numbers (cf. [146]).

(7) Clearly, the mapping of 7.5.11 fails to approximate the multiplication of \mathbb{Q}_p in contrast to the case of \mathbb{Z}_p .

Indeed, let m and n in G be presented as $m := p^{M-1}$ and $n := p^{M+1}$. Then $j(m) = p^{-1}$ and $j(n) = p$, and so $j(m)j(n) = 1$. The two cases are possible: If $2M \leq N$ then $j(mn) = p^M \approx 0$. On the other hand, if $2M > N$ then, since we discuss the addition of $\mathbb{Z}/p^N\mathbb{Z}$, we obtain $mn = p^{M-(N-M)}$. Here $mn \notin G_f$ for $N - M$ is infinite.

By analogy we may prove that the hyperapproximant of 7.5.1 to the additive group of the field \mathbb{R} fails to hyperapproximate the multiplication of \mathbb{R} for whatever Δ .

(8) Vershik and Gordon [503] proved the approximability of a nilpotent Lie group whose Lie algebra has a basis with rational structure constants and studied the class of discrete groups approximable by finite groups in detail.

The article [503] raised the question of approximability of classical simple Lie groups, in particular, $SO(3)$. The article [6] by Alekseev, Glebskiĭ, and Gordon gives a negative answer to this question by proving that for a compact Lie group G to be approximable by finite groups it is necessary and sufficient that to each $\varepsilon > 0$ there is a finite subgroup H of G serving as a ε -net for G with respect to some metric that defines the topology of G . This article [6] also contains the definition of approximability of normed commutative Hopf algebras [6] by finite-dimensional bialgebras as well as a proof that a compact group is approximable if and only if its commutative Hopf algebra is approximable by finite-dimensional commutative Hopf algebras.

7.6. Discrete Approximation of Function Spaces on a Locally Compact Abelian Group

Using the results of 7.4, we proceed with discrete approximation of the Hilbert function space over a locally compact abelian group.

In the sequel we let G stand for the primal locally compact group, while denoting the dual of G by \widehat{G} .

7.6.1. We distinguish some (left)invariant metric ρ on G , which enables us to rephrase the definition of sequential approximant 7.4.11 as follows:

A sequence $((G_n, j_n))_{n \in \mathbb{N}}$ of couples of finite groups G_n and maps $j_n : G_n \rightarrow G$ is a *sequential approximant* or *approximating sequence* to G provided that to all $\varepsilon > 0$ and all compact $K \subset G$ there is $N > 0$ such that for all $n > N$ the following hold:

- (1) $j_n(G_n)$ is an ε -net for K ;
- (2) If \circ_n is the multiplication on G_n then $\rho(j_n(g_1 \circ_n g_2^{\pm 1}), j_n(g_1)j_n(g_2)^{\pm 1}) < \varepsilon$ for all $g_1, g_2 \in j_n^{-1}(K)$;
- (3) $j_n(e_n) = e$, with e_n and e the units of G_n and G , respectively.

A locally compact group G is *approximable* provided that G possesses a sequential approximant.

Recall that if μ is the (distinguished) Haar measure on G and $((G_n, j_n))_{n \in \mathbb{N}}$ is a sequential approximant to G then every bounded μ -almost everywhere continuous function $f : G \rightarrow \mathbb{C}$ of rapid decay is integrable and

$$\int_G f d\mu = \lim_{n \rightarrow \infty} \Delta_n \sum_{g \in G_n} f(j_n(g)),$$

with $\Delta_n := \frac{\mu(U)}{|j_n^{-1}(U)|}$ (see 7.4.13 (1) which distinguishes some compact neighborhood U about the zero of U such that $\mu(U) = 1$). The numerical sequence (Δ_n) is a *sequential normalizing factor* for $((G_n, j_n))_{n \in \mathbb{N}}$. As a sequential normalizing factor we may take each sequence (Δ'_n) equivalent to (Δ_n) .

Given a compact group G , we put $\Delta_n := |G_n|^{-1}$, while we let $\Delta_n := 1$ if G is a discrete group.

We now consider a dual couple of sequential approximants $((G_n, j_n))_{n \in \mathbb{N}}$ and $((\widehat{G}_n, \widehat{j}_n))_{n \in \mathbb{N}}$ for G , cf. 7.4.14. If (Δ_n) is a sequential normalizing factor for $((G_n, j_n))_{n \in \mathbb{N}}$ (with respect to μ) then $(\widehat{\Delta}_n)$, with $\widehat{\Delta}_n := (|G_n| \Delta_n)^{-1}$, is a sequential normalizing factor for $((\widehat{G}_n, \widehat{j}_n))_{n \in \mathbb{N}}$ (with respect to the Haar measure $\widehat{\mu}$). Note that if G_n is a finite abelian group then the dual \widehat{G}_n of G_n is isomorphic to G_n and so $|G_n| = |\widehat{G}_n|$.

The Fourier transform $\mathcal{F}_n : L_2(G_n) \rightarrow L_2(\widehat{G}_n)$ acts by the rule

$$\mathcal{F}_n(\varphi)(\chi) := \Delta_n \sum_{g \in G_n} \varphi(g) \overline{\chi(g)},$$

while the inverse Fourier transform \mathcal{F}_n^{-1} takes the shape

$$\mathcal{F}_n^{-1}(\psi)(g) = \widehat{\Delta}_n \sum_{\chi \in \widehat{G}_n} \psi(\chi) \chi(g).$$

We will also denote the Fourier transform of f by \widehat{f} .

Given $p \geq 1$, we let $\mathcal{S}_p(G)$ stand for the space comprising $f : G \rightarrow \mathbb{C}$ such that $|f|^p$ is a bounded μ -almost everywhere continuous function of rapid decay. If $f \in \mathcal{S}_2(G)$ then we put $T_n f := f \circ j_n : G_n \rightarrow \mathcal{C}$ and $\widehat{T}_n \widehat{f} := \widehat{f} \circ \widehat{j}_n : \widehat{G}_n \rightarrow \mathcal{C}$. (Thus, $T_n f$ is the table of f at the knots of $j_n(G_n)$, while $\widehat{T}_n \widehat{f}$ is the table of \widehat{f} at the knots of $\widehat{j}_n(\widehat{G}_n)$.) By 7.4.15

$$\lim_{n \rightarrow \infty} \widehat{\Delta}_n \sum_{\chi \in \widehat{G}_n} |\widehat{T}_n \widehat{f}(\chi) - \mathcal{F}_n(T_n f)(\chi)|^2 = 0.$$

7.6.2. In the sequel we let $L_p(G_n)$ and $L_p(\widehat{G}_n)$ stand for the spaces \mathbb{C}^{G_n} and $\mathbb{C}^{\widehat{G}_n}$ furnished with the respective norms

$$\|\varphi\|_n^{(p)} := \left(\Delta_n \sum_{g \in G_n} |\varphi(g)|^p \right)^{1/p}, \quad \|\psi\|_n^{(p)} := \left(\widehat{\Delta}_n \sum_{\chi \in \widehat{G}_n} |\psi(\chi)|^p \right)^{1/p}.$$

In case $p = 2$ we denote them by X_n and \widehat{X}_n , omitting the subscript $p = 2$ in the symbols of their norms. By analogy, we put $X := L_2(G)$ and $\widehat{X} := L_2(\widehat{G})$. Finally, we let Y (\widehat{Y}) stand for $\mathcal{S}_2(G)$, a dense subspace of X (for $\mathcal{S}_2(\widehat{G})$, a dense subspace of \widehat{X} , respectively). We imply a distinguished couple of sequential approximants $((G_n, J_n))_{n \in \mathbb{N}}$ and $((\widehat{G}_n, \widehat{J}_n))_{n \in \mathbb{N}}$.

The sequences $((L_p(G_n), T_n))_{n \in \mathbb{N}}$ and $((L_p(\widehat{G}_n), \widehat{T}_n))_{n \in \mathbb{N}}$ are discrete approximants to $L_p(\mu)$ and $L_p(\widehat{\mu})$ respectively.

◁ Immediate from 7.4.10 and 7.4.13 (2). ▷

7.6.3. In the sequel we confine exposition to the case of a group with a compact and open subgroup. We start with a discrete group G . In this event $X = l_2(G)$ and the definition of sequential approximant simplifies significantly.

If G is a discrete group and $((G_n, J_n))_{n \in \mathbb{N}}$ is a sequential approximant to G (cf. 7.4.11 and 7.4.12) then

- (1) $\varinjlim J_n(G_n) = G$;
- (2) $(\forall a, b \in G)(\exists n_0 \in \mathcal{N})(\forall n > n_0)(\forall g, h \in G_n)$
 $(J_n(g) = a, J_n(h) = b \rightarrow J_n(g \circ_n h^{\pm 1}) = ab^{\pm 1})$;
- (3) $J_n(e_n) = e$.

Since the Haar integral of a function f on a discrete group produces the sum of the values of f , we may write the integrability condition as

$$(4) \quad (\forall f \in l_1(G))(\forall \varepsilon > 0)(\exists^{\text{fin}} K \subset G) \left(\sum_{\xi \in G-K} |f(\xi)| < \varepsilon \right).$$

From (1) it is straightforward that

$$(5) \quad (\forall^{\text{fin}} K \subset G)(\exists n_0 \in \mathcal{N})(\forall n > n_0)(J_n(G_n) \supset K).$$

A subset of a discrete group is compact whenever it is finite. Consequently, $\Delta_n = 1$ with Δ_n a normalizing factor. From (4) and (5) we readily infer that every integrable function has rapid decay. Hence, $Y = X$, and $((X_n, T_n))_{n \in \mathbb{N}}$ is a strong discrete approximant.

It is an easy matter to construct an isometric embedding $\iota_n : X_n \rightarrow X$ serving as a right inverse of T_n and satisfying the condition

$$\sup_n \sup_{\|z\|=1} (\inf\{\|x\|_n : T_n x = z\}) < +\infty.$$

Indeed, we may put

$$\iota_n(\varphi)(\xi) := \begin{cases} \varphi(g), & \xi = J_n(g), \\ 0, & \xi \notin J_n(G_n). \end{cases}$$

7.6.4. We now address the general case of a locally compact abelian group G with a compact and open subgroup K .

Put $L := G/K$. Then L is a discrete group and $\widehat{L} \subset \widehat{G}$, since $\widehat{L} = \{p \in \widehat{G} : p|_K = 1\}$. Let μ stand as before for the Haar measure on G satisfying $\mu(K) = 1$. Then the dual Haar measure $\widehat{\mu}$ on \widehat{G} enjoys the property $\widehat{\mu}(\widehat{L}) = 1$. The discrete group \widehat{K} is isomorphic to \widehat{G}/\widehat{L} . Let $\{a_l : l \in L\}$ be a selection of pairwise distinct representatives of the cosets of G/K and let $\{p_h : h \in \widehat{K}\}$ be a selection of pairwise distinct representatives of the cosets comprising \widehat{G}/\widehat{L} . Note that $p_h(k) = h(k)$ for all $k \in K$.

(1) If $a \in G$ and $p \in \widehat{G}$ then there is a unique quadruple of $l \in L$, $k \in K$, $h \in \widehat{K}$, and $s \in \widehat{L}$ such that

$$a = a_l + k, \quad p = p_h + s, \quad p(a) = p_h(a_l)s(l)h(k).$$

Take $\varphi \in L_2(G)$ and $l \in L$. Denote by φ_l a function in $L_2(K)$ such that $\varphi_l(k) = \varphi(a_l + k)$. Obviously, $\|\varphi\|^2 = \sum_{l \in L} \|\varphi_l\|^2$. Using the Fourier transform on the compact group K , we arrive at the formula $\varphi_l(k) = \sum_{h \in \widehat{K}} c_{lh}h(k)$. The correspondence $\varphi \leftrightarrow (c_{lh})_{l \in L, h \in \widehat{K}}$, which is denoted in the sequel by ι , is a unitary isomorphism between the Hilbert spaces $L_2(G)$ and $l_2(L \times \widehat{K})$, where

$$l_2(L \times \widehat{K}) := \left\{ (c_{lh})_{l \in L, h \in \widehat{K}} : \sum_{l \in L, h \in \widehat{K}} |c_{lh}|^2 < +\infty \right\}.$$

Clearly, ι depends on the particular choice of $\{a_l : l \in L\}$ for G/K .

By analogy, given $\psi \in L_2(\widehat{G})$ and $h \in \widehat{K}$, we define $\psi_h \in L_2(\widehat{L})$ by the rule $\psi_h(s) := \psi(p_h + s)$. Then $\psi_h(s) = \sum_{l \in L} d_{lh}s(l)$ and the correspondence $\psi \leftrightarrow (d_{lh})_{l \in L, h \in \widehat{K}}$, which is denoted in the sequel by $\widehat{\iota}$, is again a unitary isomorphism between the Hilbert spaces $L_2(\widehat{L})$ and $l_2(L \times \widehat{K})$. This isomorphism depends again on the particular choice of a selection $\{p_h : h \in \widehat{K}\}$ of pairwise distinct representatives of the cosets comprising \widehat{G}/\widehat{L} . Thus,

$$\psi(p_h + s) = \sum_{l \in L} d_{hl}s(l)$$

for $\psi \in L_2(\widehat{G})$.

Take $\varphi \in L_2(G)$ and insert the expression for $p(a)$ from 7.6.4 (1) in the Fourier transform

$$\widehat{\varphi}(p) = \int_G \varphi(g)\overline{p(g)} d\mu(g).$$

Then

$$\widehat{\varphi}(p_h + s) = \sum_{l \in L} \int_{\widehat{K}} \varphi_l(k) \overline{p_h(a_l) s(l) h(k)} d\mu(k).$$

Since $\int_{\widehat{K}} \varphi_l(k) \overline{h(k)} d\mu(k) = c_{lh}$ and $\overline{s(l)} = s(-l)$, infer that

$$\widehat{\varphi}(p_h + s) = \sum_{l \in L} \overline{p_h(a_{-l})} c_{-lh} s(l).$$

Comparing the last formula with the above representation for $\psi(p_h + s)$, we arrive at the following conclusion:

(2) The Fourier transform $\mathcal{F} : L_2(G) \rightarrow L_2(\widehat{G})$ amounts to the unitary operator in $l_2(L \times \widehat{K})$ with matrix

$$f(h, l, h', l') = \overline{p_{h'}(l')} \delta_{h, h'} \delta_{-l, l'}.$$

Denote by D_{test} the subspace of $L_2(G)$ comprising φ such that φ and $\widehat{\varphi}$ are both compactly supported. Assume also that \widehat{D}_{test} stands for the subspace of $L_2(\widehat{G})$ which is defined by exactly the same property. Clearly, $\mathcal{F}(D_{test}) = \widehat{D}_{test}$.

7.6.5. Suppose that $\varphi \in L_2(G)$, $\iota(f) = (c_{lh})_{l \in L, h \in \widehat{K}}$, and

$$\widetilde{\iota}(\widehat{f}) = (d_{lh})_{l \in L, h \in \widehat{K}}.$$

Then $\varphi \in D_{test}$ if and only if there are finite sets $A \subset L$ and $B \subset \widehat{K}$ such that $c_{lh} = 0$ for $(l, h) \notin A \times B$. In this event there also exist finite sets $R \subset L$ and $S \subset \widehat{K}$ such that $d_{lh} = 0$ for $(l, h) \notin R \times S$.

◁ The claim follows on observing that the matrix $f(h, l, h', l')$ has a sole nonzero entry in each row and each column. ▷

Let $((G_n, j_n))_{n \in \mathbb{N}}$ and $(\widehat{G}_n, \widehat{j}_n)_{n \in \mathbb{N}}$ be a dual couple of sequential approximants to G .

7.6.6. There is an index $n_0 \in \mathbb{N}$ such that, for all $n > n_0$, the sets $K_n := j_n^{-1}(K)$ and $\widehat{L}_n := \widehat{j}_n^{-1}(\widehat{L})$ are subgroups of G_n and \widehat{G}_n ; $((K_n, j_n|_{K_n}))_{n \in \mathbb{N}}$ and $((\widehat{L}_n, \widehat{j}_n|_{\widehat{L}_n}))_{n \in \mathbb{N}}$ are sequential approximants to K and \widehat{L} ; and \widehat{L}_n is the dual of $L_n := G_n/K_n$.

◁ Take $N \approx +\infty$. We must check that $K_N := j_N^{-1}(*K)$ is a subgroup of G_N . Assume that $j_N(a) \in *K$ and $j_N(b) \in *K$. Then $j_N(a \pm b) \approx j_N(a) \pm j_N(b) \in *K$, implying that $j_N(a \pm b) \in *K$, since K is a compact and open subgroup. Therefore, K_N is a subgroup.

Further, $(K_N, j_N|_{K_N})$ obviously meets the conditions 7.4.1 (1, 2), which proves that $((K_n, j_n|_{K_n}))_{n \in \mathbb{N}}$ is a sequential approximant to K ; cf. 7.4.11. The claim about the dual sequential approximant is established by analogy.

We are left with demonstrating that \widehat{L}_N is the dual of $L_N := G_N/K_N$, which amounts to the equivalence

$$\widehat{j}_N(\kappa) \in {}^* \widehat{L} \leftrightarrow \kappa|_{K_N} \equiv 1$$

holding for all $\kappa \in \widehat{G}_N$.

If $\widehat{j}_N(\kappa) \in {}^* \widehat{L}$ then $\widehat{j}_N(\kappa)|_{*K} \equiv 1$. Hence, $1 = \widehat{j}_N(\kappa)(j_N(g)) \approx \kappa(g)$ for $g \in K_N$. Note that κ and g are nearstandard since \widehat{L} and K are compact. Consequently, $\kappa|_{K_N} \approx 1$ and $\kappa|_{K_N} \equiv 1$ by 7.2.11 (1).

Assume conversely that $\kappa|_{K_N} \equiv 1$. Then $\kappa(g) = 1$ for all $g \in G_N$ satisfying $j_N(g) \approx 0$. This, along the lines of the proof of 7.4.10, implies that $\widehat{j}_N(\kappa) \in \text{nst}({}^* \widehat{G})$. Let $k \in {}^* K$. Then there is some g in K_N satisfying $j_N(g) \approx k$. Therefore, $\widehat{j}_N(\kappa)(k) \approx \widehat{j}_N(\kappa)(j_N(g)) \approx \kappa(g) = 1$, and so $j_N(\kappa)|_{*K} \approx 1$. Since \widehat{K} is a discrete group, conclude that $j_N(\kappa)|_{*K} \equiv 1$. \triangleright

7.6.7. In what follows, K_n and \widehat{L}_n are subgroups of G_n and \widehat{G}_n for all $n \in \mathbb{N}$.

A sequential approximant $((L_n, j'_n))_{n \in \mathbb{N}}$ of a discrete group L is *compatible* with a sequential approximant $((G_n, j_n))_{n \in \mathbb{N}}$ provided that to each finite set $B \subset L$ there is an index $n_0 \in \mathbb{N}$ satisfying the condition: $j'_n(\lambda) = l \in B$ implies $j_n^{-1}(l) = \lambda$ for all $n > n_0$ and $\lambda \in L_n$.

We may rephrase this as follows:

A sequential approximant $((L_n, j'_n))_{n \in \mathbb{N}}$ to a discrete group L is compatible with a sequential approximant $((G_n, j_n))_{n \in \mathbb{N}}$ if and only if, given $N \approx +\infty$ and a standard $l \in L$, we have $j'_N(\lambda) = l \leftrightarrow j_N^{-1}(l) = \lambda$ for all $\lambda \in L_N$.

A sequential approximant $((\widehat{G}_n, \widehat{j}_n))_{n \in \mathbb{N}}$ to \widehat{G} is a dual to a sequential approximant $((G_n, j_n))_{n \in \mathbb{N}}$ to G if and only if the following hold for all $N \approx +\infty$:

- (1) If $\chi \in \widehat{G}_N$ is such that $\chi(g) \approx 1$ for all $g \in j_N^{-1}(\text{nst}({}^* G))$ then $\widehat{j}_N(\chi) \approx 0$;
- (2) If $j_N(g) \in \text{nst}({}^* G)$ and $\widehat{j}_N(\chi) \in \text{nst}({}^* \widehat{G})$ then $\widehat{j}_N(\chi)(j_N(g)) \approx \chi(g)$.

\triangleleft This is immediate from 7.4.6 and 7.4.11. \triangleright

7.6.8. A dual sequential approximant $((L_n, j'_n))_{n \in \mathbb{N}}$ to some sequential approximant $((\widehat{L}_n, \widehat{j}_n|_{\widehat{L}_n}))_{n \in \mathbb{N}}$ to \widehat{L} is compatible with $((G_n, j_n))_{n \in \mathbb{N}}$.

A dual sequential approximant $((\widehat{K}_n, \widehat{j}_n))_{n \in \mathbb{N}}$, with $\widehat{K}_n := \widehat{G}_n/\widehat{L}_n$, to a sequential approximant $((K_n, j_n|_{K_n}))_{n \in \mathbb{N}}$ is compatible with $((\widehat{G}_n, \widehat{j}_n))_{n \in \mathbb{N}}$.

◁ Take $\lambda \in L_N$ and $j'_N(\lambda) = l \in L$. Note that $\widehat{j}_N(\kappa)(l) \approx \kappa(\lambda)$ for $\kappa \in \widehat{L}_N$ by the definition of dual sequential approximant in 7.6.7. We have to prove that $j_N^{-1}(l) = \lambda$. To this end, observe that there is a sole element $\lambda' \in G_N$ satisfying $j_N(\lambda') = l$. Indeed, if $j_N(\lambda') = j_N(\lambda'') = l$ then $j_N(a - b) \approx j_N(a) - j_N(b) \in {}^*K$ for $a \in \lambda'$ and $b \in \lambda''$. Since K is compact and open; therefore, $j_N(a - b) \in K$, and so $a - b \in K_N$ and $\lambda' = \lambda''$. Clearly, $\kappa(\lambda) \approx \kappa(\lambda')$ for all $\kappa \in \widehat{L}_N$ and, arguing as in the proof of Theorem 7.4.10, we arrive at the equality $\lambda' = \lambda$. ▷

7.6.9. Let $\{a_l : l \in L\}$ be a selection of pairwise distinct representatives of the cosets comprising G/K . Assume given $\varepsilon > 0$ and a finite set $B \subset L$. Then for all sufficiently large $n \in \mathbb{N}$ there is a selection $\{\alpha_\lambda : \lambda \in L_n\}$ of pairwise distinct representatives of the cosets comprising G_n/K_n such that $\rho(a_{j'_n(\lambda)}, \alpha_\lambda) < \varepsilon$ for all $\lambda \in j_n^{-1}(B)$. Here $((L_n, j'_n))_{n \in \mathbb{N}}$ is a sequential approximant to L compatible with $((G_n, j_n))_{n \in \mathbb{N}}$.

◁ Clearly, the claim may be rephrased nonstandardly as follows:

Let $\{a_l : l \in L\}$ be a selection of pairwise distinct representatives of the cosets comprising G/K . Then to each $N \approx +\infty$ there is a selection $\{\alpha_\lambda : \lambda \in L_N\}$ of pairwise distinct representatives of the cosets comprising G_N/K_N such that $j_N(\alpha_\lambda) \approx a_l$ for all $\lambda \in j_N^{-1}(L)$, with $j'_N(\lambda) = l$.

Put $R := j_N^{-1}(\{a_l : l \in L\}) \subset G_N$. Define the internal equivalence \sim on R by the rule $g \sim h \leftrightarrow g - h \in K_N$ and assign $R' := R/\sim$. Also, define the internal set $S \subset R'$ by putting $S := \{r \in R' : |r| = 1\}$, and assign $S' := \bigcup S$. Then $j_N^{-1}(a_l) \in S'$ for every standard $l \in L$.

Put $S'' := \{\lambda \in L_N : (\exists g \in S')(g \in \lambda)\}$ and let T stand for a selection of pairwise distinct representatives of the cosets belonging to $L - S''$. It is then clear that $S' \cap T = \emptyset$ and $S' \cup T$ is a selection of pairwise distinct representatives of the cosets belonging to L_N . ▷

7.6.10. EXAMPLE. Let G be the additive group of \mathbb{Q}_p , the field of p -adic numbers.

Choose two sequences of integers $r, s \rightarrow \infty$ and put $n := r + s$. Assume further that G_n is the additive group of the ring $\mathbb{Z}/p^n\mathbb{Z} := \{0, 1, \dots, p^n - 1\}$ (this particular presentation of this ring is material for what follows). Define $j_n : G_n \rightarrow \mathbb{Q}_p$ by the rule $j_n(k) := \frac{k}{p^r}$. Then $((G_n, j_n))_{n \in \mathbb{N}}$ is a sequential approximant to G ; cf. 7.5.9.

The dual group $\widehat{G} := \widehat{\mathbb{Q}_p}$ is isomorphic with \mathbb{Q}_p . An arbitrary character of \mathbb{Q}_p takes the shape $\chi_\xi(\eta) := \exp(2\pi i\{\xi\eta\})$ for $\xi \in \mathbb{Q}_p$, with $\xi \rightarrow \chi_\xi$ a topological isomorphism. We may thus identify \mathbb{Q}_p with $\widehat{\mathbb{Q}_p}$ and consider a dual sequential approximant as another sequential approximant to \mathbb{Q}_p . We also identify G_n with \widehat{G}_n . Then the dual sequential approximant $((G_n, \widehat{j}_n))_{n \in \mathbb{N}}$ is given by the formula $\widehat{j}_n(m) := \frac{m}{p^s}$; cf. 7.5.10.

As a compact and open subgroup K of G we take the additive group of \mathbb{Z}_p , the ring of p -adic integers. Then $K_n := j_n^{-1}(K) = p^r G_n := \{kp^r : k := 0, 1, \dots, p^s - 1\}$.

To define the quotient group $L := G/K$, we denote by $\mathbb{Q}^{(p)}$ the additive group of the rationals of the shape $\frac{m}{p^l}$, with $l > 0$. Clearly, L is isomorphic to $\mathbb{Q}^{(p)}/\mathbb{Z}$. The quotient group $L_n := G_n/K_n$ is isomorphic to $\mathbb{Z}/p^r\mathbb{Z} := \{0, 1, \dots, p^r - 1\}$. Define the embedding $j'_n : L_n \rightarrow L$ by the rule $j'_n(t) := \frac{t}{p^r}$. It is an easy matter to check that $((L_n, j'_n))_{n \in \mathbb{N}}$ is a sequential approximant to L compatible with the sequential approximant $((G_n, j_n))_{n \in \mathbb{N}}$.

The sets $\{\frac{k}{p^l} : k < p^l, k|p\}$ and $\{0, 1, \dots, p^r - 1\}$ are complete families of pairwise distinct representatives of the cosets comprising G/K and $L_n := G_n/K_n$ respectively, which meet the hypothesis of 7.6.9. We omit the simple proof of this fact.

In accord with the above identifications, we have $\widehat{L} = \mathbb{Z}_p = K$ and $\widehat{K} = L$. Therefore, $\widehat{L}_n = \{0, 1, \dots, p^r - 1\} \simeq \{kp^s : k = 0, 1, \dots, p^r - 1\}$, $\widehat{K}_n = \{0, 1, \dots, p^s - 1\}$; and if $\widehat{j}'_n : \widehat{K}_n \rightarrow L$ is given by the rule $\widehat{j}'_n(u) := \frac{u}{p^s}$ then $((\widehat{K}_n, \widehat{j}'_n))_{n \in \mathbb{N}}$ is a sequential approximant to $\widehat{G}/\widehat{L} := \widehat{K} = L$ compatible with the sequential approximant $((G_n, \widehat{j}_n))_{n \in \mathbb{N}}$.

The set $\{0, 1, \dots, p^s - 1\}$ is a selection of pairwise distinct representative of the cosets comprising $\widehat{K}_n := G_n/\widehat{L}_n$, which meets the hypothesis of 7.6.9 with respect to this sequential approximant.

(1) *The matrix of the Fourier transform (in the context of 7.6.4 (2)) is as follows:*

$$f((m, l), (u, v), (m', l'), (u', v')) = \exp\left(-\frac{2\pi i m' u'}{p^{l'+v'}}\right) \delta_{(m,l),(m',l')} \delta_{(p^v-u,v),(u',v')}.$$

◁ Put $\Gamma_p := \{(m, l) : m|p, 0 \leq m < p^l\}$. On easy grounds, we may identify $l_2(L \times \widehat{K})$ with $l_2(\Gamma_p^2)$. Since the equality $\frac{u}{p^v} = \frac{p^v-u}{p^v}$ holds in $\mathbb{Q}^{(p)}$, we complete the proof. ▷

Analogous considerations apply to the finite Fourier transform $\mathcal{F}_n : L_2(G_n) \rightarrow L_2(\widehat{G}_n)$. More precisely,

(2) *The matrix of the finite Fourier transform \mathcal{F}_n (in the context of 7.6.4 (2)) is as follows:*

$$f(l, v, l', v') = \exp\left(-\frac{2\pi i l' v'}{p^n}\right) \delta_{p^r-l,l'} \delta_{vv'}.$$

◁ Identify $L_2(G_n)$ with $l_2(L_n \times \widehat{K}_n)$ by the rule

$$\varphi(l + kp^r) := \sum_{h=0}^{p^s-1} c_{lh} \exp \frac{2\pi i k h}{p^s}$$

$(\varphi \in L_2(G_n), 0 \leq l < p^r - 1, 0 \leq k < p^s - 1).$

Also, identify $L_2(\widehat{G}_n)$ with $l_2(L_n \times \widehat{K}_n)$ by the rule

$$\psi(v + tp^s) := \sum_{l=0}^{p^r-1} d_{lv} \exp \frac{2\pi i t l}{p^r}$$

$(\psi \in L_2(\widehat{G}_n), 0 \leq v < p^s - 1, 0 \leq t < p^r - 1).$

We thus come to the chain of equalities:

$$\begin{aligned} \mathcal{F}_n(\varphi) &= \widehat{\varphi}(v + tp^s) = \frac{1}{p^s} \sum_{u=0}^{p^n-1} \varphi(u) \exp \left(-\frac{2\pi i u (v + tp^s)}{p^n} \right) \\ &= \frac{1}{p^s} \sum_{l=0}^{p^r-1} \sum_{k=0}^{p^s-1} \sum_{h=0}^{p^s-1} c_{lh} \exp \left(\frac{2\pi i k h}{p^s} \right) \exp \left(-\frac{2\pi i (l + kp^r)(v + tp^s)}{p^n} \right) \\ &= \sum_{l=0}^{p^r-1} c_{lv} \exp \left(-\frac{2\pi i l v}{p^n} \right) \exp \left(-\frac{2\pi i l t}{p^r} \right), \end{aligned}$$

completing the proof. ▷

7.6.11. If $((X_n, T_n))_{n \in \mathbb{N}}$ is a discrete approximant then it usually fails to be strong. However, we may slightly change $((X_n, T_n))_{n \in \mathbb{N}}$ to make it strong. In this subsection we construct a strong discrete approximant $((X_n, S_n))_{n \in \mathbb{N}}$ to a space X which enjoys 6.2.6 (1) as well as the property $\|T_n f - S_n f\|_n \rightarrow 0$ holding for all f belonging to some dense subset of Y . Clearly, in this event the strong discrete approximant with S_n determines the same embedding $\mathbf{t} : X \rightarrow \mathcal{X}$ as the discrete approximant with T_n .

Some strong discrete approximant in the case of \mathbb{R}^n was constructed in [86]. Here we consider only the case of a group with a compact and open subgroup. As was mentioned in 7.6.3, if G is a discrete group then every discrete approximant $((X_n, T_n))_{n \in \mathbb{N}}$ is strong, with 6.2.6 (1) holding.

(1) Let G be a compact group. Then the normalizing factor takes the shape $\Delta_n := |G_n|^{-1}$. In this event the discrete approximant $((X_n, T_n))_{n \in \mathbb{N}}$ may fail to be strong. The dense subspace $Y \subseteq X$ consists of bounded almost everywhere continuous functions and it is easy to check that T_n is not extendible to the whole of X in general. We define $S_n : X \rightarrow X_n$ by the rule:

$$S_n(f)(g) := \sum_{\chi \in \widehat{G}_n} \widehat{f}(\widehat{j}_n(\chi)) \chi(g).$$

(2) Assume now that G is a locally compact abelian group with a compact and open subgroup K . Assume further that $L := G/K$, while $((G_n, j_n))_{n \in \mathbb{N}}$ and $((\widehat{G}_n, \widehat{j}_n))_{n \in \mathbb{N}}$ is a dual couple of sequential approximants to G . Suppose that K_n meets the hypothesis of 7.6.6, $L_n := G_n/K_n$, and $((L_n, j'_n))_{n \in \mathbb{N}}$ is a sequential approximant to L compatible with $((G_n, j_n))_{n \in \mathbb{N}}$.

Distinguish a selection $\{a_l : l \in L\}$ of pairwise distinct representatives of the cosets comprising L . Dropping to a subsequence, if need be, we may assume by 7.6.9 that to each $n \in \mathbb{N}$ there is a selection $\{\alpha_\lambda : \lambda \in L_n\}$ of pairwise distinct representatives of the cosets comprising L_n such that

$$\lim_{n \rightarrow \infty} j_n(\alpha_{j_n^{-1}(l)}) = a_l \quad (l \in L).$$

Denote by S'_n the same operator from $L_2(K)$ to $L_2(K_n)$ as in (1). We also define $S_n : X \rightarrow X_n$ as

$$S_n \varphi(\alpha_\lambda + \xi) := S'_n \varphi_{j'_n(\lambda)}(\xi) \quad (\xi \in K_n, \lambda \in L_n).$$

Here, as well as above, $\varphi_l(k) := \varphi(a_l + k)$.

7.6.12. *In case the group under study is compact, $((X_n, S_n))_{n \in \mathbb{N}}$ is a strong discrete approximant to X such that $\|T_n f - S_n f\|_n \rightarrow 0$ for all $f \in Y$ and 6.2.6(1) holds.*

◁ Since $\{\chi(g) : \chi \in \widehat{G}_n\}$ is an orthonormal basis for X_n ; therefore, $\|S_n(f)\|^2 = \sum_{\chi \in \widehat{G}_n} |\widehat{f}(\widehat{j}_n(\chi))|^2$. The group \widehat{G}_n is discrete and so the discrete approximant $((\widehat{X}_n, \widehat{T}_n))_{n \in \mathbb{N}}$ to \widehat{X} will be strong. Hence,

$$\lim_{n \rightarrow \infty} \sum_{\chi \in \widehat{G}_n} |\widehat{f}(\widehat{j}_n(\chi))|^2 = \|\widehat{f}\|^2 = \|f\|^2,$$

and so, $\|S_n(f)\| \rightarrow \|f\|$.

If $f \in Y$ then, on applying Theorem 7.4.15 to the inverse Fourier transform, infer that $\|T_n(f) - \mathcal{F}_n^{-1}\widehat{T}_n\widehat{f}\| \rightarrow 0$ as $n \rightarrow \infty$. Now, by the definition of the inverse Fourier transform $\mathcal{F}_n^{-1}\widehat{T}_n\widehat{f} = S_n(f)$.

To demonstrate that 6.2.6 (1) holds for the discrete approximant $((X_n, S_n))_{n \in \mathbb{N}}$, define the embedding $\iota_n : X_n \rightarrow X$ by the rule

$$\iota_n(\varphi)(\xi) := \sum_{\chi \in \widehat{G}_n} \mathcal{F}_n(\varphi)(\chi) \widehat{j}_n(\chi)(\xi).$$

Then $\iota_n(X_n) = \{\sum_{\chi \in \widehat{G}_n} c_\chi \widehat{j}_n(\chi)\}$. Moreover, it is easy that $\iota_n(S_n(f)) = f$ for all $f \in \iota_n(X_n)$. Therefore, if $p_n : X \rightarrow \iota_n(X_n)$ is an orthoprojection then $S_n = \iota_n^{-1} \circ p_n$, yielding 6.2.6 (1). \triangleright

7.6.13. *In case we study a locally compact group with a compact and open subgroup, $((X_n, S_n))_{n \in \mathbb{N}}$ is a strong discrete approximant to X satisfying 6.2.6.(1). Moreover, $\|T_n\varphi - S_n\varphi\|_n \rightarrow 0$ as $n \rightarrow \infty$ for all $\varphi \in \mathcal{S}_2(G)$.*

\triangleleft As shown above, the correspondence $\varphi \leftrightarrow \{\varphi_l : l \in L\}$ implements a unitary isomorphism between the Hilbert spaces X and $\prod_{l \in L} X^{(l)}$, where each $X^{(l)}$ coincides with $L_2(K)$. By analogy, the correspondence $\psi \leftrightarrow \{\psi_\lambda : \lambda \in L_n\}$, with $\psi_\lambda(\xi) := \psi(\alpha_\lambda + \xi)$, implements a unitary isomorphism between the Hilbert spaces X_n and $\prod_{l \in L} X_n^{(\lambda)}$, where each $X_n^{(\lambda)}$ coincides with $L_2(K_n)$. Identifying unitarily isomorphic Hilbert spaces, obtain

$$S_n(\{\varphi_l : l \in L\}) = \{S'_n \varphi_{j_n(\lambda)} : \lambda \in L_n\}.$$

This yields the first part of the claim since S'_n satisfies 7.6.12.

To prove the second part, we first assume that φ is a compactly supported continuous function. Then there is a standard finite set $A \subset L$ such that $\varphi(a_l + k) = 0$ for all $k \in K$ whenever $l \notin A$. Take $N \approx +\infty$. It suffices to show only that $\|T_N\varphi - S_N\varphi\| \approx 0$. Let $\{\alpha_\lambda : \lambda \in L_N\}$ be a selection of pairwise distinct representatives of the cosets comprising L_N such that the nonstandard version of 7.6.9 holds. Given $g \in K_N$, we then see that $T_N\varphi(\alpha_\lambda + g) = \varphi \circ j_N(\alpha_\lambda + g) \neq 0$ if and only if $j'_N(\lambda) \in A$. If $j'_N(\lambda) = l \in A$ then, considering the definition of hyperapproximant in 7.4.1 (2) and the relation $j_N(\alpha_\lambda) \approx a_l$, we may write $T_N\varphi(\alpha_\lambda + g) = \varphi(j_N(\alpha_\lambda + g)) \approx \varphi(j_N(\alpha_\lambda) + j_N(g)) \approx \varphi(a_l + j_M(g)) = T_N\varphi_l(g)$. We also use here the fact that a compactly supported continuous function φ is uniformly continuous, and so from $\alpha \approx \beta$ it follows that $\varphi(\alpha) \approx \varphi(\beta)$ even for nonstandard α and β ; cf. 2.3.12. Now, the definition of S_n in 7.6.11 (2) yields $S_N\varphi(\alpha_\lambda + g) = S'_N\varphi_l(g)$. If $j'_N(\lambda) = l \in A$ then $T_N\varphi_l \approx S'_N\varphi_l$ by 7.6.12, and if $j'_N(\lambda) = l \notin A$ then $\varphi_l = 0$ and so $S'_N\varphi_l = 0$. Since the size of A is standardly finite; therefore, $\|T_N\varphi - S_N\varphi\|_N \approx 0$.

Let φ be an arbitrary member of $\mathcal{S}_2(G)$. Distinguish an arbitrary standard real $\varepsilon > 0$. Then there is a compactly supported continuous function ψ satisfying $\|\varphi - \psi\| < \varepsilon$. By the definition of discrete approximation, $\|T_N(\varphi) - T_N(\psi)\|_N = \|T_N(\varphi - \psi)\|_N \approx \|\varphi - \psi\|$. By the same reason, $\|S_N(\varphi) - S_N(\psi)\|_N \approx \|\varphi - \psi\|$, implying that $\|T_N(\varphi) - T_N(\psi)\|_N + \|S_N(\varphi) - S_N(\psi)\|_N < 2\varepsilon$. Moreover, $\|T_N\varphi - S_N\varphi\|_N \leq \|T_N(\varphi) - T_N(\psi)\|_N + \|T_N(\psi) - S_N(\psi)\|_N + \|S_N(\varphi) - S_N(\psi)\|_N < 5\varepsilon$ because $\|T_N(\psi) - S_N(\psi)\|_N \approx 0$. Since $\varepsilon > 0$ is an arbitrary standard real; therefore, $\|T_N\varphi - S_N\varphi\|_N \approx 0$. \triangleright

By analogy we may define $\widehat{S}_n : L_2(\widehat{G}) \rightarrow L_2(G_n)$ satisfying 7.6.13. Assume that $\{\pi_\nu : \nu \in \widehat{K}_n\}$ is a selection of pairwise distinct representatives of the cosets comprising $\widehat{G}_n/\widehat{L}_n$ such that 7.6.9 holds for the sequential approximant $((\widehat{G}_n, \widehat{J}_n))_{n \in \mathbb{N}}$, and $\widehat{S}'_n : L_2(\widehat{L}) \rightarrow L_2(\widehat{L}_n)$ is an operator satisfying 7.6.11 for the sequential approximant $((\widehat{L}_n, \widehat{J}_n|_{\widehat{L}_n}))_{n \in \mathbb{N}}$ to \widehat{L} . Then $\widehat{S}_n\psi(\pi_\nu + \eta) = \widehat{S}'_n\psi_{\widehat{J}'_n(\nu)}(\eta)$ for all $\nu \in \widehat{K}_n$ and $\eta \in \widehat{L}_n$.

7.6.14. We now return to 7.6.10. If $\varphi \in L_2(\mathbb{Q}_p)$ is an almost everywhere continuous function then we have the chain of equalities

$$\begin{aligned} (T_n\varphi)(j + kp^r) &= \varphi(j_n(j + kp^r)) = \varphi\left(\frac{j}{p^r} + k\right) \\ &= \varphi_{(l,m)}(k) = \sum_{(u,v) \in \Gamma_p} c_{(l,m)(u,v)} \exp\left(\frac{2\pi iku}{p^v}\right), \end{aligned}$$

with $(l, m) \in \Gamma_p$ and $l/p^m = j/p^r$.

To calculate $S_n\varphi$, note that $L = \widehat{\mathbb{Z}}_p$ and so $\widehat{J}'_n : \widehat{K}_n \rightarrow L$ is a dual approximant of $J_n|_{K_n}$. Using 7.6.11 (1), obtain

$$\begin{aligned} S_n\varphi(j + kp^r) &= S'_n\varphi_{\widehat{J}'_n(j)}(j_n(kp^r)) = S'_n\varphi_{(l,m)}(k) \\ &= \sum_{w \in \widehat{K}_n} c_{(l,m)\widehat{J}'_n(w)} \exp\frac{2\pi iw k}{p^s} = \sum_{v \leq s} \sum_{(u,v) \in \Gamma_p} c_{(l,m)(u,v)} \exp\left(\frac{2\pi iku}{p^v}\right). \end{aligned}$$

If $\varphi \in D_{test}$ (cf. 7.6.5) while r and s satisfy the relations $m > r, v > s, c_{(l,m)(u,v)} = 0$, and $n = r + s$, then the above expressions for S_n and T_n yield $S_n\varphi = T_n\varphi$.

Comparing 7.6.10 (1) and 7.6.10 (2) with the expression for S_n , infer that $\widehat{S}_n(\widehat{f}) = F_n(S_n f)$ for all $f \in L_2(\mathbb{Q}_p)$. Moreover, if $f \in D_{test}$ then $\widehat{T}_n(\widehat{f}) = F_n(T_n f)$. Since D_{test} is dense in X , we have established Theorem 7.4.15 in the case under consideration.

7.6.15. Given $N \approx +\infty$, we define the space $\mathcal{X} := X_N^\#$ and the operator $\mathbf{t} : X \rightarrow \mathcal{X}$ as in 6.2.3. Here $X_N := L_2(G_N)$ and $X := L_2(G)$ (cf. 7.6.2). Proposition 6.2.4 shows that it is important to know the condition for $\varphi \in X_N^{(b)}$ to satisfy the containment $\varphi^\# \in \mathbf{t}X$.

We say that $\varphi \in X_N^{(b)}$ is *proxy-standard*; in symbols, $\varphi \in \text{proxy}(X_N^{(b)})$ provided that there is some $f \in X$ satisfying $\varphi^\# = \mathbf{t}(f)$. We will use the notations:

$$\begin{aligned} H(G_N) &:= G_N - j_N^{-1}(\text{proxy}(*G)), & H(\widehat{G}_N) &:= G_N - \widehat{j}_N^{-1}(\text{proxy}(*\widehat{G})), \\ H(L_N) &:= L_N - j_N'^{-1}(L), & H(\widehat{K}_N) &:= \widehat{K}_N - \widehat{j}_N'^{-1}(\widehat{K}). \end{aligned}$$

Theorem. An element $\varphi \in X_N^{(b)}$ is proxy-standard if and only if the following hold:

- (1) $\Delta_N \sum_{g \in B} |\varphi(g)|^2 \approx 0$ for every internal subset B of $H(G_N)$;
- (2) $\widehat{\Delta}_N \sum_{\chi \in C} |\mathcal{F}_N(\varphi)(\chi)|^2 \approx 0$ for every internal subset C of $H(\widehat{G}_N)$.

$\triangleleft \rightarrow$: Assume that $\varphi := \mathbf{t}(f)$. Since D_{test} is dense in X , we may suppose that to each standard $\varepsilon > 0$ there is some ψ in D_{test} satisfying

$$\Delta_N \sum_{g \in G_N} |\varphi(g) - \psi(j_N(g))|^2 < \varepsilon.$$

The Fourier transform is an isometry, and $\widehat{\psi} \circ \widehat{j}_N \approx \mathcal{F}_N(\psi \circ j_N)$ by Theorem 7.4.10. Hence,

$$\widehat{\Delta}_N \sum_{\chi \in \widehat{G}_N} |\mathcal{F}_N(\varphi)(\chi) - \widehat{\psi}(\widehat{j}_N(\chi))|^2 < \varepsilon.$$

Clearly, the same estimates hold in the case when summation ranges over some internal subsets of G_N and \widehat{G}_N . The functions ψ and $\widehat{\psi}$ are compactly supported, and so $*\text{supp } \psi \subset \text{proxy}(*G)$ and $*\text{supp } \widehat{\psi} \subset \text{proxy}(*\widehat{G})$. Consequently, given arbitrary internal sets $B \subset H(G_N)$ and $C \subset H(\widehat{G}_N)$, we have $\psi \circ j_N|_B = 0$ and $\widehat{\psi} \circ j_N|_C = 0$. The above estimates now yield

$$\Delta_N \sum_{g \in B} |\varphi(g)|^2 < \varepsilon, \quad \widehat{\Delta}_N \sum_{\chi \in C} |\mathcal{F}_N(\varphi)(\chi)|^2 < \varepsilon.$$

Since $\varepsilon > 0$ is an arbitrary standard real, the proof of necessity is complete.

\leftarrow : Assume now that φ enjoys the conditions (1) and (2). Distinguish some complete families $\{\alpha_\lambda : \lambda \in L_N\}$ and $\{\pi_\nu : \nu \in \widehat{K}_N\}$ of pairwise distinct representatives of the cosets comprising L_N and $\widehat{G}_N/\widehat{L}_N$ respectively, which satisfy the

nonstandard version of 7.6.9 (cf. the proofs of 7.6.8 and 7.6.9). If $\lambda \in L_N$ and $k \in K_N$ then

$$\varphi(\alpha_\lambda + k) = \sum_{\nu \in \widehat{K}_N} \sigma_{\lambda, \nu} \nu(k).$$

To complete the prove, we need two auxiliary facts.

(A) If $P \subset H(L_N)$ and $Q \subset H(\widehat{K}_N)$ are internal then

$$\sum_{\lambda \in P} \sum_{\nu \in \widehat{K}_N} |\sigma_{\lambda \nu}|^2 \approx 0, \quad \sum_{\nu \in Q} \sum_{\lambda \in L_N} |\sigma_{\lambda, \nu}|^2 \approx 0.$$

◁ Note in this case that the normalizing factors take the shape $\Delta_N := |K_N|^{-1}$ and $\widehat{\Delta}_N := |L_N|^{-1}$; cf. 7.6.1 (as K we take some relatively compact open neighborhood about the zero of G). Let $P \subset H(L_N)$ be an internal set. Then $B = P + K_N \subset H(G_N)$. Since $\{\nu(k) : \nu \in K_N\}$ is an orthonormal basis for $L_2(K_N)$, infer from (A) that:

$$\begin{aligned} 0 &\approx |K_N|^{-1} \sum_{g \in B} |\varphi(g)|^2 = |K_N|^{-1} \sum_{\lambda \in P} \sum_{k \in K_N} |\varphi(\alpha_\lambda + k)|^2 \\ &= |K_N|^{-1} \sum_{\lambda \in P} \sum_{k \in K_N} \left| \sum_{\nu \in \widehat{K}_N} \sigma_{\lambda \nu} \nu(k) \right|^2 = \sum_{\lambda \in P} \sum_{\nu \in \widehat{K}_N} |\sigma_{\lambda \nu}|^2. \end{aligned}$$

Similarly, given $\nu \in \widehat{K}_N$ and $\gamma \in \widehat{L}_N$, note that

$$\mathcal{F}_N(\varphi)(\pi_\nu + \gamma) = \sum_{\lambda \in \widehat{L}_N} \widehat{\sigma}_{\lambda, \nu} \gamma(\lambda).$$

Hence,

$$\sum_{\nu \in Q} \sum_{\lambda \in L_N} |\widehat{\sigma}_{\lambda, \nu}|^2 \approx 0$$

for each internal subset Q of $H(\widehat{K}_N)$.

Repeating the calculations that led to the formula of 7.6.4(2) for the Fourier transform \mathcal{F}_N , obtain $\widehat{\sigma}_{\lambda, \nu} = \pi_\nu(\alpha_\lambda) \sigma_{-\lambda \nu}$. Hence, $|\widehat{\sigma}_{\lambda, \nu}| = |\sigma_{-\lambda \nu}|$, yielding the second of the equalities in question. ▷

The groups L and K are countable since G is separable. Therefore, there are increasing sequences of finite subsets $A'_m \subset L$ and $B'_m \subset \widehat{K}$ such that $L = \bigcup_{m \in \mathbb{N}} A'_m$ and $\widehat{K} = \bigcup_{m \in \mathbb{N}} B'_m$. Put $A_m := j'^{-1}_N(A'_m)$ and $B_m := \widehat{j}'^{-1}_N(B'_m)$. Define the sequence $(\varphi_m)_{m \in \mathbb{N}} \subset X_N$ by the rule

$$\varphi_m(\alpha_\lambda + k) := \begin{cases} \sum_{\nu \in B_m} \sigma_{\lambda \nu} \nu(k), & \text{if } \lambda \in A_m, \\ 0, & \text{if } \lambda \notin A_m. \end{cases}$$

(B) The following holds in \mathcal{X} :

$$\varphi^\# = \lim_{m \rightarrow \infty} \varphi_m^\#.$$

\triangleleft It suffices to check that to each standard $\varepsilon > 0$ there is a standard natural m_0 satisfying $\|\varphi - \varphi_m\| < \varepsilon$ for all $m > m_0$. In much the same way as this is done for $L_2(G)$, we may easily demonstrate that the correspondence $\varphi \leftrightarrow \{\sigma_{\lambda,\nu} : \lambda \in L_N, \nu \in \widehat{K}_N\}$ implements a unitary isomorphism between the Hilbert spaces $L_2(G_N)$ and $l_2(L_N \times \widehat{K}_N)$. Therefore,

$$\|\varphi - \varphi_m\|^2 = \sum_{(\lambda,\nu) \in L_N \times \widehat{K}_N - A_m \times B_m} |\sigma_{\lambda\nu}|^2.$$

It is clear that $L \subset A'_M \subset {}^*L$ and $\widehat{K} \subset B'_M \subset {}^*\widehat{K}$ for every $M \approx +\infty$. Also, $P := L_N - A_M \subset H(L_N)$ and $Q := \widehat{K}_N - B_M \subset H(\widehat{K}_N)$. If $S \subset L_N \times \widehat{K}_N - A_M \times B_M$ then

$$\sum_{(\lambda,\nu) \in S} |\sigma_{\lambda\nu}|^2 \leq \sum_{(\lambda,\nu) \in P \times \widehat{K}_N} |\sigma_{\lambda\nu}|^2 + \sum_{(\lambda,\nu) \in L_N \times Q} |\sigma_{\lambda\nu}|^2.$$

The two sums on the right side of the above equality are infinitesimal by (A), and so $\|\varphi - \varphi_M\|^2 \approx 0$. Consider the internal set $C := \{m \in {}^*\mathbb{N} : \|\varphi - \varphi_m\| < \varepsilon\}$. As we have just established, C contains all infinite hypernaturals M . By underflow, there is some natural m_0 such that C contains all $m > m_0$, which ends the proof. \triangleright

Since $\|\varphi\|$ is limited, $\sum_{\lambda \in \widehat{L}_N} |\sigma_{\lambda\nu}|^2$ is limited too, implying that so is each of the hyperreals $\sigma_{\lambda\nu}$. Given $l \in L$ and $h \in \widehat{K}$, put $c_{lh} := {}^\circ\sigma_{j_N^{-1}(l)j_N^{-1}(k)}$ and define $f_m \in D_{test}$ for $m \in \mathbb{N}$ by the rule

$$f_m(a_l + \xi) := \begin{cases} \sum_{h \in B'_m} c_{lh} h(\xi), & \text{if } l \in A'_m, \\ 0, & \text{if } l \notin A'_m. \end{cases}$$

Using 7.6.9 and 7.6.12, infer that

$$S_N(f_m)(\alpha_\lambda + k) = \begin{cases} \sum_{\nu \in B_m} c_{j_N(\lambda)j_N(\nu)} \nu(k), & \text{if } \lambda \in A_m, \\ 0, & \text{if } \lambda \notin A_m, \end{cases}$$

and $S_N(f_m) \approx T_N(f_m)$. Therefore, $S_N(f_m)^\# = \mathbf{t}(f_m)$. At the same time,

$$\|S_N(f_m) - \varphi_m\|^2 = \sum_{\lambda \in A_m} \sum_{\nu \in B_m} |\sigma_{\lambda\nu} - c_{j_N(\lambda)j_N(\nu)}|^2.$$

This is an infinitesimal since $\sigma_{\lambda\nu} \approx c_{j_N(\lambda)j_N(\nu)}$ and the size of $A_m \times B_m$ is standardly finite. From (B) it follows now that $\varphi^\# = \lim_{m \rightarrow \infty} \mathbf{t}(f_m)$, yielding $\varphi^\# \in \mathbf{t}(X)$. \triangleright

7.6.16. COMMENTS.

(1) The results of this section are taken from the article [5] by Albeverio, Gordon, and Khrennikov.

(2) As regards 7.6.10 (2), it is worth noting that the celebrated Cooley–Tukey algorithm for the fast Fourier transform rests on exactly the same calculations; cf. [25].

(3) Inspecting 7.6.10, we observe that the analogous considerations apply to each of the groups $\mathbb{Q}_{\mathbf{a}}$, where $\mathbf{a} := (a_n)_{n \in \mathbb{Z}}$ is an arbitrary sequence of naturals (see the definitions in [174] wherein these groups are denoted by $\Omega_{\mathbf{a}}$). A couple of dual sequential approximants to $\mathbb{Q}_{\mathbf{a}}$ is described in the Introduction to the book [146]. It is well known (for instance, see [151]) that each totally disconnected locally compact abelian group is isomorphic to $\mathbb{Q}_{\mathbf{a}}$ with a suitable \mathbf{a} .

(4) Theorem 7.4.15, together with the construction of a strong discrete approximant and Proposition 7.6.12, yields Theorem 6.1 in [85] of approximation of locally compact abelian groups by finite groups in the sense of Weyl systems in the case of a group with a compact and open subgroup. To derive Theorem 6.1 of [85] from Theorem 7.4.15 in the case of \mathbb{R}^n , it is necessary to use the strong approximant of $L_2(\mathbb{R}^n)$ that is given in [86].

7.7. Hyperapproximation of Pseudodifferential Operators

In this section we hyperapproximate pseudodifferential operators on a locally compact abelian group with a compact open subgroup.

7.7.1. Let G be a locally compact abelian group and let \widehat{G} be the dual of G .

(1) Given a sufficiently good function $f : G \times \widehat{G} \rightarrow \mathbb{C}$, we may define a (possible unbounded) linear operator $A_f : L_2(G) \rightarrow L_2(G)$ by the rule

$$A_f \psi(x) := \int_{\widehat{G}} f(x, \chi) \widehat{\psi}(\chi) \chi(x) d\widehat{\mu}(\chi) \quad (\psi \in L_2(G)).$$

We say that A_f is a *pseudodifferential operator with symbol f* .

(2) An n th approximant $A_f^{(n)} : \mathbb{C}^{(G_n)} \rightarrow \mathbb{C}^{(G_n)}$ to A_f is defined as

$$A_f^{(n)} \varphi(x) := \widehat{\Delta}_n \sum_{\chi \in \widehat{G}_n} f(j_n(x), \widehat{j}_n(\chi)) \mathcal{F}_n(\varphi)(\chi) \chi(x) \quad (\varphi \in X_n, x \in G_n).$$

We confine exposition to the case of G a group with a compact and open subgroup K . As before, $L := G/K$, and we distinguish some selections $\{a_l : l \in L\}$ and $\{p_h : h \in \widehat{K}\}$ of pairwise distinct representatives of the cosets comprising L and $\widehat{K} := \widehat{G}/\widehat{L}$.

(3) Clearly, the mappings $(j_n, \widehat{j}_n) : G_n \times \widehat{G}_n \rightarrow G \times \widehat{G}$ define a sequential approximant to $G \times \widehat{G}$. Denote by $S_n^{(2)} : L_2(G_n \times \widehat{G}_n) \rightarrow L_2(G \times \widehat{G})$ the mapping that is defined in 7.6.13 for this approximant. Therefore, $((L_2(G_n \times \widehat{G}_n), S_n^{(2)}))_{n \in \mathbb{N}}$ is a strong discrete approximant to $L_2(G \times \widehat{G})$ and we may define another approximant for A_f by the rule:

$$B_f^{(n)} \varphi(x) := \widehat{\Delta}_n \sum_{\chi \in \widehat{G}_n} (S_n^{(2)} f)(x, \chi) \mathcal{F}_n(\varphi)(\chi) \chi(x).$$

7.7.2. For A_f to be a Hilbert–Schmidt operator it is necessary and sufficient that $f \in L_2(G \times \widehat{G})$. In this event,

$$\|A_f\| \leq \iint_{G \times \widehat{G}} |f(x, \chi)|^2 d\mu \otimes \widehat{\mu}(x, \chi).$$

◁ The claim, almost evident as stated, is well known in the classical theory of pseudodifferential operators (for instance, see [33]).

Indeed, the proof is immediate on observing that the kernel of A_f has the shape $K(x, y) := (\mathcal{F}_{\widehat{G}} f)(x, x - y)$, with $\mathcal{F}_{\widehat{G}}$ the Fourier transform in the second variable. Hence, $\|K(x, y)\|_{L_2(G^2)} = \|f(x, \chi)\|_{L_2(G \times \widehat{G})}$. ▷

7.7.3. Theorem. If A_f is a Hilbert–Schmidt operator then

(1) The sequence $(B_f^{(n)})$ in 7.7.1 (3) is uniformly bounded, i.e.,

$$(\exists n_0)(\forall n > n_0) \left(\|B_n\| \leq \|f(x, \chi)\|_{L_2(G \times \widehat{G})} \right);$$

(2) The sequence $(B_f^{(n)})$ converges discretely to A_f with respect to the strong discrete approximant $((X_n, S_n))_{n \in \mathbb{N}}$ in 7.6.13; moreover, this convergence is uniform;

(3) If $f \in \mathcal{S}_2(G \times \widehat{G})$ (with the notation of 7.6.1) then $\|A_n - B_n\| \rightarrow 0$. Moreover, (1) and (2) hold for $(A_f^{(n)})$.

◁ The proof proceeds in a few steps.

(a): Take $N \approx +\infty$. Clearly, $\|B_n\| \leq \|S_n^{(2)} f\|_n$ for all $n \in {}^*\mathbb{N}$. We recall here in particular that $\|\mathcal{F}_n\| = 1$ and $|\chi(x)| = 1$. By the definition of discrete approximant, $\|S_N^{(2)} f\|_N \approx \|f\|$. If $f \in \mathcal{S}_2(G \times \widehat{G})$ then $\|S_N^{(2)} f - T_N f\|_N \approx 0$. This yields (1) and (3), and so we are left with (2).

(b): Let ι and $\widehat{\iota}$ stand for the unitary isomorphisms of 7.6.4 (1) between $L_2(G)$ and $l_2(L \times \widehat{K})$ and between $L_2(\widehat{G})$ and $l_2(L \times \widehat{K})$. Then A_f , as an endomorphism of $l_2(L \times \widehat{K})$, takes the shape

$$(\iota A_f \varphi)(l, h) = \sum_{h', l'} (\widetilde{\iota f})(l, h', l + l', h - h') p_{h'}(a_l) (\widehat{\iota \varphi})(l', h').$$

◁ Indeed, given $\varphi \in L_2(G)$, find

$$\begin{aligned} \widehat{\varphi}(p_h s) &= \sum_{l \in L} (\widehat{\iota \varphi})(l, h) s(l), \\ A_f \varphi(a_l + k) &= \sum_{h \in \widehat{K}} (\iota A_f \varphi)(l, h) h(k). \end{aligned}$$

Similarly,

$$f(a_l + k, p_h + s) = \sum_{h' \in \widehat{K}} \sum_{l' \in L} (\widetilde{\iota f})(l, h, l', h') s(l') h'(k),$$

with $\widetilde{\iota f} := (\iota \otimes \widehat{\iota}) f$. Simple calculations, resting on 7.6.4 (1) and 7.7.1 (1), yield the claim. ▷

(c): Distinguish some selections of pairwise distinct representatives $\{\alpha_\lambda : \lambda \in L_N\}$ and $\{\pi_\nu : \nu \in \widehat{K}_N\}$ of the cosets comprising L_N and \widehat{K}_N so as to satisfy the nonstandard version of 7.6.9 (cf. the proof of 7.6.9). Then

$$\begin{aligned} (\iota B_f^{(N)} \varphi)(\lambda, \nu) &= \sum_{\nu', \lambda'} (\widetilde{\iota f})(j'_N(\lambda), \widehat{j}'_N(\nu'), j'_N(\lambda + \lambda'), \widehat{j}'_N(\nu - \nu')) \\ &\quad \times p_{\nu'}(\alpha_\lambda) (\widehat{\iota}_N F_N(\varphi))(\lambda', \nu'). \end{aligned}$$

◁ Indeed, it is obvious that if $K \times \widehat{L}$ is a compact and open subgroup of $G \times \widehat{G}$. Moreover, $L \times \widehat{K} = G \times \widehat{G} / K \times \widehat{L}$ and $\{(a_l, p_h) : l \in L, h \in \widehat{K}\}$ is a selection of pairwise distinct representatives of the cosets comprising L and \widehat{K} . Furthermore, $\{(\alpha_\lambda, \pi_\nu) : \lambda \in L_N, \nu \in \widehat{K}_N\}$ is a selection of pairwise distinct representatives of the cosets comprising $L_N \times \widehat{K}_N = G_N \times \widehat{G}_N / K_N \times \widehat{L}_N$ which satisfies the nonstandard version of 7.6.9. It is easy to see then that

$$\begin{aligned} &S_N^{(2)} f(\alpha_\lambda + \xi, \pi_\nu + \eta) \\ &= \sum_{\lambda' \in L_N} \sum_{\nu' \in \widehat{K}_N} (\widetilde{\iota f})(j'_N(\lambda), \widehat{j}'_N(\nu), j'_N(\lambda') \widehat{j}'_N(\nu')) \overline{\eta(\lambda')} \nu'(\xi), \end{aligned}$$

where $((L_N, j'_N))_{n \in \mathbb{N}}$ is some sequential approximant to L compatible with the sequential approximant $((G_N, j_N))_{n \in \mathbb{N}}$ (see the definition in 7.6.7). Moreover, $((\widehat{K}_N, \widehat{j}'_N))_{n \in \mathbb{N}}$ is a sequential approximant to \widehat{K} dual to the sequential approximant $((K_N, j_N|_{K_N}))_{n \in \mathbb{N}}$ to K .

As above, the unitary isomorphisms $\iota_N : L_2(G_N) \rightarrow l_2(L_N \times \widehat{K}_N)$ and $\widehat{\iota}_N : L_2(\widehat{G}_N) \rightarrow l_2(L_N \times \widehat{K}_N)$ are defined by the rules

$$\begin{aligned}\varphi(\alpha_\lambda + \xi) &:= \sum_{\nu \in \widehat{K}_N} (\iota_N \varphi)(\lambda, \nu) \nu(\xi), \\ \psi(\pi_\nu + \eta) &:= \sum_{\lambda \in L_N} (\widehat{\iota}_N \psi)(\lambda, \nu) \eta(\lambda).\end{aligned}$$

Arguing as in (b), complete the proof. \triangleright

(d): If $\psi \in L_2(G)$ then $\|\iota_N B_f^{(N)} S_N \psi - \iota_N S_N A_f \psi\| \approx 0$.

\triangleleft We start with auxiliary calculations. It is immediate by definition that

$$(d_1) \quad (\iota_N S_N \psi)(\lambda, \nu) = (\psi)(j'_N(\lambda), \widehat{j}'_N(\nu)).$$

By 7.6.4 (2),

$$(d_2) \quad \widehat{\iota} \psi(j'_N(\lambda), \widehat{j}'_N(\nu)) = \psi(-j'_N(la), \widehat{j}'_N(\nu)) \overline{p_{\widehat{j}'_N(\nu)}(a_{-j'_N(\lambda)})}.$$

Analogous calculations, together with (d₁), yield

$$(d_3) \quad (\widehat{\iota}_N \mathcal{F}_N(S_N \psi))(\lambda, \nu) = (\psi)(j'_N(-\lambda), \widehat{j}'_N(\nu)) \overline{\pi_\nu(\alpha_{-\lambda})}.$$

From (b), (d₁), (c), and (d₃) we infer the two formulas:

$$(d_4) \quad \begin{aligned} \iota_N S_N A_f \psi(\lambda, \nu) &= \sum_{l', h'} (\widehat{\iota} f)(j'_N(\lambda), h', j'_N(\lambda) \\ &\quad + l', \widehat{j}'_N(\nu) - h') p_{h'}(a_{j'_N(\lambda)}) \widehat{\iota} \psi(l', h'); \end{aligned}$$

$$(d_5) \quad \begin{aligned} (\iota B_f^{(N)} S_N \psi)(\lambda, \nu) &= \sum_{\nu', \lambda'} (\widehat{\iota} f)(j'_N(\lambda), \widehat{j}'_N(\nu'), j'_N(\lambda + \lambda'), \widehat{j}'_N(\nu - \nu')) \pi_{\nu'}(\alpha_\lambda) \\ &\quad (\psi)(j'_N(-\lambda'), \widehat{j}'_N(\nu')) \overline{\pi_{\nu'}(\alpha_{-\lambda'})}. \end{aligned}$$

Assume now that $f \in D_{test}^{(2)}$. This implies that there are some finite sets $A \subset L$ and $B \subset \widehat{K}$ such that $(\widehat{\iota} f)(l, h, l', h') = 0$ whenever $(l, h, l', h') \notin (A \times B)^2$.

The space $D_{test}^{(2)}$ is dense in $L_2(G \times \widehat{G})$. Choose standard finite sets A_1 and B_1 so that $(A - A) \cup A \subset A_1 \subset L$ and $(B - B) \cup B \subset B_1 \subset \widehat{K}$. Put $C := j'^{-1}_N(A_1)$ and $D := \widehat{j}'^{-1}_N(B_1)$.

From (d₄) it is then clear that $(\iota_N S_N A_f \psi)(\lambda, \nu) = 0$ for $(\lambda, \nu) \notin C \times D$. Hence we must confine the domains of l and h' to A_1 and B_1 respectively. Since the finite sets A_1 and B_1 are standard, from 7.6.3 it follows that $j'_N(\alpha \pm \alpha') = j'_N(\alpha) \pm j'_N(\alpha')$ and $\widehat{j}'_N(\beta \pm \beta') = \widehat{j}'_N(\beta) \pm \widehat{j}'_N(\beta')$ for $\alpha, \alpha' \in C$ and $\beta, \beta' \in D$.

These observations, together with (b), show that we may rewrite (d₄) as follows:

$$\begin{aligned}
 (d_6) \quad & (\iota_N S_N A_f \psi)(\lambda, \nu) \\
 &= \sum_{\lambda' \in C} \sum_{\nu' \in D} (\widehat{\iota f})(j'_N(\lambda), \widehat{j}'_N(\nu'), j'_N(\lambda + \lambda'), \widehat{j}'_N(\nu - \nu')) \\
 & \quad \times (\iota \psi)(-j'_N(\lambda), \widehat{j}'_N(\nu)) \overline{p_{\widehat{j}'_N(\nu')} (a_{j'_N(\lambda)}) p_{\widehat{j}'_N(\nu')} (a_{-j'_N(\lambda')})}.
 \end{aligned}$$

By the same reason, we may assume that the variables λ' and ν' in the sum on the right side of (d_5) range over C and D respectively, while $(\iota B_f^{(N)} S_N \psi)(\lambda, \nu) = 0$ for $(\lambda, \nu) \notin C \times D$.

Comparing (d_4) and (d_6) , note that the terms under the summation sign differ in the coefficients $\pi_{\nu'}(\alpha_\lambda) \pi_{\nu'}(\alpha_{-\lambda'})$ in (d_5) and

$$\overline{p_{\widehat{j}'_N(\nu')} (a_{j'_N(\lambda)}) p_{\widehat{j}'_N(\nu')} (a_{-j'_N(\lambda')})}$$

in (d_6) . However, these coefficients are infinitely close to one another by the non-standard version of 7.6.9. Consequently, the left sides of (d_5) and (d_6) are infinitely close to one another since the sums on the right sides have standardly many nonzero terms. This implies (d) , and 7.7.3(2) for $f \in D_{test}^{(2)}$. \triangleright

The general case results from the following auxiliary proposition.

(e): *If 7.7.3(2) holds for A_f , with f ranging over some dense subset Y of $L_2(G \times \widehat{G})$, then 7.7.3(2) holds for all $f \in L_2(G \times \widehat{G})$.*

\triangleleft It suffices to prove that $\|S_N A_f - B_f^{(N)} S_N\| \approx 0$. Take an arbitrary standard $\varepsilon > 0$ and choose $\psi \in Y$ so that $\|f - \psi\| < \varepsilon$. Then

$$\begin{aligned}
 \|S_N A_f - B_f^{(N)} S_N\| &\leq \|S_N A_f - S_N A_\psi\| \\
 &+ \|S_N A_\psi - B_\psi^{(N)} S_N\| + \|B_\psi^{(N)} S_N - B_f^{(N)} S_N\|.
 \end{aligned}$$

Note further that $\|S_N A_\psi - B_\psi^{(N)} S_N\| \approx 0$ by hypothesis. By 6.2.2, $\|S_N\|$ is limited, and so

$$\begin{aligned}
 \|S_N A_f - S_N A_\psi\| &\leq \circ \|S_N\| \|A_{f-\psi}\| \leq \circ \|S_N\| \|f - \psi\| \leq \circ \|S_N\| \varepsilon, \\
 \|B_\psi^{(N)} S_N - B_f^{(N)} S_N\| &\leq \|B_{\psi-f}^{(N)}\| \circ \|S_N\| \leq \circ \|S_N\| \|f - \psi\| \leq \circ \|S_N\| \varepsilon.
 \end{aligned}$$

The arbitrary choice of $\varepsilon > 0$ completes the proof of (e). \triangleright

We have happily completed the proof of Theorem 7.7.3 \triangleright

7.7.4. *If A_f is a Hilbert–Schmidt hermitian operator then the following hold:*

- (1) *The spectrum $\sigma(A_f)$ coincides with the set of nonisolated limit points of $\bigcup_n \sigma(B_f^{(n)})$;*
- (2) *If $0 \neq \lambda \in \sigma(A_f)$ and J is a neighborhood about λ containing no points of $\sigma(A_f)$ but λ then λ is the only nonisolated limit point of $J \cap \bigcup_n \sigma(B_f^{(n)})$.*

\triangleleft Immediate from 7.7.3 and 6.2.8, on recalling 6.2.3(1) and 7.6.13. \triangleright

7.7.5. The operators $A_f^{(n)}$ and $B_f^{(n)}$ may fail to be hermitian even in the case when A_f is selfadjoint. This compels us in this context to formulate the rest of Theorem 6.2.8 as follows:

In the context of 7.7.4, assume additionally that A_f is selfadjoint and the sequence $(C^{(n)})$, with $C^{(n)}$ a hermitian endomorphism of X_n , possesses the property: $\|B_f^{(n)} - C^{(n)}\|_n \rightarrow 0$ as $n \rightarrow \infty$. Then the following hold:

(1) If

$$M_n^\lambda := \sum_{\nu \in \sigma(C^{(n)}) \cap J} C^{(n)(\nu)}$$

(cf. 7.7.4(2)) then

$$\dim(M_n^\lambda) = \dim(A_f^{(\lambda)}) = s$$

for all sufficiently large n , and there is a sequence of orthonormal bases $(f_1^n, \dots, f_s^n)_{n \in \mathbb{N}}$ for M_n^λ converging discretely to some orthonormal basis (f_1, \dots, f_s) for $A_f^{(\lambda)}$ with respect to the discrete approximant $((X_n, T_n))_{n \in \mathbb{N}}$;

(2) In the context of (1), if $f_1, \dots, f_s \in \mathcal{S}_2(G)$ then the sequence of orthonormal bases $((f_1^n, \dots, f_s^n))_{n \in \mathbb{N}}$ converges discretely to the orthogonal basis (f_1, \dots, f_s) with respect to the discrete approximant $((X_n, T_n))_{n \in \mathbb{N}}$.

7.7.6. We now address the so-called Schrödinger-type operators.

Let $f : G \times \widehat{G} \rightarrow \mathbf{C}$ take the shape

$$f(x, \chi) := a(x) + b(\chi) \quad (x \in G, \chi \in \widehat{G}).$$

If $a(x) \geq 0$ and $a(x) \rightarrow \infty$ as $x \rightarrow \infty$ then A_f is a *Schrödinger-type operator*.

Up to the end of this section, we assume that $a(\cdot)$ and $b(\cdot)$ are almost everywhere continuous locally bounded real functions on G and \widehat{G} respectively, with $a(x) \rightarrow \infty$ as $x \rightarrow \infty$ and $b(\chi) \rightarrow \infty$ as $\chi \rightarrow \infty$. We also suppose that G and the sequential approximant to G under study satisfy all hypotheses of 7.7.5.

It is easy to see that the definition of A_f in 7.7.1 (1) leads in this context to

$$A_f \psi(x) = a(x) \psi(x) + \overset{\vee}{b} * \psi(x),$$

where $*$ is the convolution on $L_1(G)$ and $\overset{\vee}{b}$ is the inverse Fourier transform of the function b treated as a distribution on \widehat{G} .

Similarly, the discrete approximant of 7.7.1 (2) now satisfies the formula

$$A_f^{(n)}\varphi(x) = a(j_n(x))\varphi(x) + \mathcal{F}_n^{(-1)}(b \circ \widehat{j}_n) * \varphi(x).$$

Under the assumptions we made, each Schrödinger-type operator on a locally compact abelian group with compact and open subgroup has discrete spectrum consisting of real eigenvalues of finite multiplicity:

$$\alpha_1 \leq \lambda_2 \leq \dots \leq \lambda_k \dots \rightarrow \infty; \quad k \rightarrow \infty.$$

The proof may be supplied along the lines of [506], wherein this was proven for \mathbb{Q}_p . It is also easy to see that the operators $A_f^{(n)}$ are selfadjoint in this environment.

7.7.7. Theorem. *Assume that A_f is a Schrödinger-type operator such that a and b satisfy the conditions of 7.7.6 and the approximation domain of A_f by the sequence $(A_f^{(n)})$ of 7.7.1 (2) is an essential domain of A_f . Then the following hold:*

- (1) *The spectrum $\sigma(A_f)$ comprises the nonisolated limit points of the set $\bigcup_n \sigma(A_f^{(n)})$;*
- (2) *If J is a neighborhood of λ containing no points $\sigma(A_f)$ but λ then λ is the only nonisolated limit point of $J \cap \bigcup_n \sigma(A_f^{(n)})$;*
- (3) *In the context of (2), if*

$$M_n^\lambda := \sum_{\nu \in \sigma(A_f^{(n)}) \cap J} A_f^{(n)(\nu)}$$

then $\dim(M_n^\lambda) = \dim(A_f^{(\lambda)}) = s$ for all sufficiently large n and there is a sequence of orthonormal bases $((f_1^n, \dots, f_s^n))_{n \in \mathbb{N}}$ for M_n^λ converging discretely to an orthonormal basis (f_1, \dots, f_s) for $A_f^{(\lambda)}$ with respect to the discrete approximant $((X_n, S_n))_{n \in \mathbb{N}}$ (and also with respect to the discrete approximant $((X_n, T_n))_{n \in \mathbb{N}}$ provided that the eigenfunctions of A_f belong to $\mathcal{S}(G)$).

◁ Without loss of generality, we may assume a and b positive functions. Then A_f and $A_f^{(n)}$ are positive operators satisfying $-1 \notin \text{cl}(\sigma(A_f) \cup \bigcup_n \sigma(A_f^{(n)}))$. By 6.2.10 it suffices to prove that $(R_{-1}(A_f^{(n)}))$ is a quasicompact sequence.

So, given $N \approx +\infty$ and $\psi \in X_n$, we will prove that if $\|(A_f^{(N)} + I)\psi\|_N$ is limited then ψ satisfies the conditions of Theorem 7.6.15. Since $\|R_{-1}(A_f^{(N)})\| \leq 1$; therefore, $\|\psi\|_N$ is limited, implying that $((A_f^{(N)} + I)\psi, \psi)$ is limited too. However,

$$\begin{aligned} ((A_f^{(N)} + I)\psi, \psi) &= (a \cdot \psi, \psi) + (\mathcal{F}_N^{(-1)}(b) * \psi, \psi) + (\psi, \psi), \\ (\mathcal{F}_N^{(-1)}(b) * \psi, \psi) &= (b \cdot \mathcal{F}_N(\psi), \mathcal{F}_N(\psi)), \end{aligned}$$

and so $(a\psi, \psi)$ and $(bF_N(\psi), F_N(\psi))$ are limited hyperreals. Assume now by way of contradiction that the first condition of Theorem 7.6.15 fails. Then there are an internal set $B \subset H(G_N)$ and a standard hyperreal $c > 0$ satisfying the inequality $\Delta_N \sum_{x \in B} |\psi(x)|^2 > c$. Since $a(x) \rightarrow \infty$ as $x \rightarrow \infty$, we may find $L \approx \infty$ satisfying $a(x) > L$ for all $x \in B$. The above implies that

$$(a\psi, \psi) \geq \Delta_N \sum_{x \in B} a(x) |\psi(x)|^2 \geq Lc,$$

which contradicts the limitedness of $(a\psi, \psi)$.

The second condition of Theorem 7.6.15 is checked by analogy. \triangleright

We now describe a class of Schrödinger-type operators satisfying the conditions of Theorem 7.7.7.

We will use the unitary isomorphisms ι and $\widehat{\iota}$ defined in the proof of Theorem 7.7.3. Despite the fact that $a \notin L_2(G)$ and $b \notin L_2(\widehat{G})$, we denote by ιa and $\widehat{\iota} b$ the functions in $l_2(L \times \widehat{K})$ such that

$$\begin{aligned} a(a_l + k) &= \sum_{h \in \widehat{K}} (\iota a)(l, h) h(k); \\ b(p_h + s) &= \sum_{l \in L} (\widehat{\iota} b)(l, h) s(l). \end{aligned}$$

7.7.8. *Let A_f be a Schrödinger-type operator. Assume also that a and b meet the above requirements and, moreover, $S(l) := \{h \in \widehat{K} : (\iota a)(l, h) \neq 0\}$ and $T(h) := \{l \in L : (\widehat{\iota} b)(l, h) \neq 0\}$ are finite sets for all $l \in L$ and $h \in \widehat{K}$. Then D_{test} is an essential domain of A_f serving as the approximation domain of A_f by $(A_f^{(n)})$.*

\triangleleft Recall that the space D_{test} comprises the functions $\varphi \in L_2(G)$ such that φ and $\widehat{\varphi}$ are compactly supported. This implies that there are finite sets $A[\varphi] \subset L$ and $B[\varphi] \subset \widehat{K}$ satisfying $(\iota\varphi)(l, h) = 0$ for $(l, h) \notin A[\varphi] \times B[\varphi]$.

We proceed by simple calculation resting on 7.6.4 (2), which yields

$$\begin{aligned} (1) \quad (\iota A_f \varphi)(l, h) &= \sum_{h' \in B[\varphi]} (\iota a)(l, h' - h) (\iota\varphi)(l, h') \\ &+ \sum_{l' \in A[\varphi]} (\widehat{\iota} b)(l - l', h) (\iota\varphi)(l', h) p_h(a_{-l}) \overline{p_h(a_{l'})}. \end{aligned}$$

This formula shows that $A_f \varphi \in D_{test}$ and, moreover,

$$\begin{aligned} A[A_f \varphi] &\subset A[\varphi] \cup \left(A[\varphi] + \bigcup_{h \in B[\varphi]} T(h) \right) = A'[\varphi]; \\ B[A_f \varphi] &\subset B[\varphi] \cup \left(B[\varphi] - \bigcup_{l \in A[\varphi]} S(l) \right) = B'[\varphi]. \end{aligned}$$

If $\varphi \in D(A_f)$ then $\iota A_f \varphi$ satisfies (1) with $A[\varphi] = L$ and $B[\varphi] = \widehat{K}$.

Given finite sets $A \subset L$ and $B \subset \widehat{K}$, denote by $P(A, B)$ the orthoprojection from $L_2(G)$ to the subspace of functions φ satisfying $A[\varphi] \subset A$ and $B[\varphi] \subset B$. Then, it is easy from (1) that $P(A, B)A_f \varphi = A_f P(A', B')\varphi$, with

$$A' := A \cup \left(A - \bigcup_{h \in B} T(h) \right), \quad B' := B \cup \left(B + \bigcup_{l \in A} S(l) \right).$$

Given $\varepsilon > 0$, we may find finite sets $A \subset L$ and $B \subset \widehat{K}$ such that $\|P(A, B)\varphi - \varphi\| < \varepsilon$ and $\|P(A, B)A_f \varphi - A_f \varphi\| < \varepsilon$. Since $A \subset A'$ and $B \subset B'$; therefore, $\|P(A', B')\varphi - \varphi\| < \varepsilon$. Hence, if $\psi := P(A', B')\varphi$ then $\psi \in D_{test}$ and, moreover, $\|\varphi - \psi\| < \varepsilon$ and $\|A_f \varphi - A_f \psi\| < \varepsilon$. Thus, D_{test} is an essential domain of A_f .

We are left with proving that $S_N(A_f \psi) \approx A_f^{(N)} S_N \psi$ for all $N \approx +\infty$ and each standard $\psi \in D_{test}$. To this end, we conveniently rewrite the definitions of A_f and $A_f^{(n)}$ in 7.7.6 as follows:

$$\begin{aligned} A_f \psi(x) &= a(x)\psi(x) + \mathcal{F}^{-1}(b\widehat{\psi})(x), \\ A_f^{(N)} \varphi(x) &= a(j_N(x))\varphi(x) + \mathcal{F}_N^{-1}(b\mathcal{F}_N \varphi)(x). \end{aligned}$$

The supports of ψ and $\widehat{\psi}$ enjoy the obvious relations:

$$\text{supp } \psi \subset \bigcup_{l \in A[\psi]} a_l + K = C, \quad \text{supp } \widehat{\psi} \subset \bigcup_{h \in B[\psi]} p_h + \widehat{L} = D.$$

From 7.7.3 $(d_1), (d_2)$ it follows that

$$\begin{aligned} (2) \quad \text{supp } S_N \psi &\subset \bigcup \{ \alpha_\lambda + K_N : \lambda \in j_N^{-1}(A[\psi]) \} = C_N, \\ \text{supp } \mathcal{F}_N S_N \psi &\subset \bigcup \{ \pi_\nu + \widehat{L}_N : \nu \in \widehat{j}_N^{-1}(B[\psi]) \} = D_N. \end{aligned}$$

However, $S_N A_f \psi = S_N(a\psi) + S_N \mathcal{F}^{-1}(b\widehat{\psi})$ and

$$A_f^{(N)} S_N \psi = a \circ j_N \cdot S_N \psi + \mathcal{F}_N^{-1}(b \circ \widehat{j}_N \cdot \mathcal{F}_N S_N \psi).$$

Since ψ is compactly supported; therefore, $a \cdot \psi$ belongs to the subspace Y of 7.6.12, and so $S_N(a \cdot \psi) \approx T_N(a \cdot \psi) = a \circ j_N \cdot \psi \circ j_N$. Since a is bounded on C , we further infer that

$$\begin{aligned} &\|a \circ j_N \cdot S_N \psi - a \circ j_N \cdot \psi \circ j_N\|_N^2 \\ &= \Delta_N \sum_{x \in C_N} |a(j_N(x))(S_N \psi(x) - \psi(j_N(x)))|^2 \approx 0. \end{aligned}$$

Consequently, $S_N(a \cdot \psi) \approx a \circ j_N \cdot S_N\psi$, and we are left with proving only that $S_N\mathcal{F}^{-1}(b \cdot \widehat{\psi}) \approx \mathcal{F}_N^{-1}(b \circ \widehat{j}_N \cdot \mathcal{F}_N S_N\psi)$.

Since $A_f\psi \in D_{test} \subset Y$ and, as we have just established, $a \cdot \psi \in Y$; therefore, $\mathcal{F}^{-1}(b \cdot \widehat{\psi}) \in Y$, implying that $S_N\mathcal{F}^{-1}(b \cdot \widehat{\psi}) \approx T_N\mathcal{F}^{-1}(b \cdot \widehat{\psi}) = \mathcal{F}^{-1}(b \cdot \widehat{\psi}) \circ j_N$. It is also clear that $\mathcal{F}_N S_N\psi \approx \mathcal{F}_N T_N\psi \approx \widehat{\psi} \circ \widehat{j}_N$ since $S_N\psi \approx T_N\psi$ and \mathcal{F}_N is bounded.

Note that $\text{supp } \widehat{\psi} \circ \widehat{j}_N \subset D_N$. Using (2) and the boundedness of b on D , we finally obtain $b \circ \widehat{j}_N \cdot \mathcal{F}_N S_N\psi \approx b \circ \widehat{j}_N \cdot \widehat{\psi} \circ \widehat{j}_N$, which implies that $\mathcal{F}_N^{-1}(b \circ \widehat{j}_N \cdot \mathcal{F}_N S_N\psi) \approx \mathcal{F}_N^{-1}(b \circ \widehat{j}_N \cdot \widehat{\psi} \circ \widehat{j}_N) \approx \mathcal{F}^{-1}(b \cdot \widehat{\psi}) \circ j_N$. \triangleright

7.7.9. We now return to the example of 7.6.10, where $G := \mathbb{Q}_p$ and $\widehat{G} := \mathbb{Q}_p$ to within isomorphism. The article [506] treats the Schrödinger-type operator with symbol

$$f(x, \chi) := a(|x|_p) + b(|\chi|_p).$$

If $a(|x|_p) \rightarrow \infty$ as $x \rightarrow \infty$ and $b(|\chi|_p) \rightarrow \infty$ as $\chi \rightarrow \infty$ then such an operator satisfies the conditions of 7.7.8 since a and b are constant functions on the cosets of $\mathbb{Q}_p/\mathbb{Z}_p$, implying that the sets $S(l)$ and $T(h)$ in 7.7.8 are singletons.

Using the dual couple of sequential approximants to \mathbb{Q}_p which is described in 7.6.10, we may easily write out the n th approximant $A_f^{(n)}$ defined in 7.7.6. Namely, given $n := r + s$, we have

$$A_f^{(n)}\varphi(k) = a(p^r |k|_p) + \frac{1}{p^n} \sum_{j,m=0}^{n-1} b(p^s |m|_p) \varphi(k-j) \exp \frac{2\pi i j m}{n}.$$

7.7.10. The concept of Weyl symbol may be abstracted to the case of a locally compact abelian group G if G admits division by 2. This means that to each $x \in G$ there is some $y \in G$ satisfying $y + y = x$ and from $y + y = 0$ it follows that $y = 0$ for all $y \in G$. In this event we denote y by $\frac{1}{2}y := y/2$ and assume that the mapping $x \mapsto \frac{1}{2}x$ is continuous on G . Note that if G admits division by 2 then so does \widehat{G} ; i.e., $\frac{1}{2}\chi(x) := \chi(\frac{1}{2}x)$.

The operator W_f with Weyl symbol $f : G \times \widehat{G} \rightarrow \mathbb{C}$ is defined by the rule

$$W_f := \iint_{G \times \widehat{G}} \widetilde{f}(y, \gamma) U_y V_\gamma \overline{\gamma y/2} d\mu \otimes \widehat{\mu}(y, \gamma),$$

with

$$\widetilde{f}(y, \gamma) := \iint_{G \times \widehat{G}} f(x, \chi) \overline{\chi}(\gamma) \overline{\gamma}(x) d\mu \otimes \widehat{\mu}(x, \chi).$$

Clearly, W_f is a symmetric operator if and only if f is a real function. If f has the shape $a(x) + b(\chi)$ then $A_f = W_f$.

It is an easy matter to calculate the kernel $K_f(x, y)$ of W_f . To this end, denote by $f^{(2)}(x, y)$ the inverse Fourier transform of $f(\cdot, \cdot)$ with respect to the second variable; i.e., $f^{(2)}(x, y) := \int_{\widehat{G}} f(x, \chi)\chi(y)d\widehat{\mu}(\chi)$. Then $K_f(x, y) = f^{(2)}(\frac{x+y}{2}, y - x)$.

If the entries G_n of the sequential approximant $((G_n, J_n))_{n \in \mathbb{N}}$ also admit division by 2 then we will say that the sequential approximant $((G_n, J_n))_{n \in \mathbb{N}}$ approximates the division by 2 on G whenever

$$(\forall \varepsilon > 0)(\forall K \in \mathcal{K})(\exists N > 0)(\forall n > N)(\forall g \in J_N^{(-1)}(K)) \\ \rho(J_n(g/2), J_n(x)/2) < \varepsilon.$$

It is easy to see that if p is an odd prime then the sequential approximant of 7.6.10 approximates the division by 2 on \mathbb{Q}_p .

If the division by 2 is approximable then we may define some sequential approximant $(W_f^{(n)})$ as follows: Define $f_n : G_n \times \widehat{G}_n \rightarrow \mathbb{C}$ as

$$f_n(g, \kappa) := f(J_n(g), \widehat{j}_n(\kappa)),$$

and let \widehat{f}_n stand for the finite Fourier transform of f_n :

$$\widehat{f}_n(h, \chi) := \frac{1}{|G_n|} \sum_{g, \kappa} f_n(g, \kappa) \overline{\chi(g)\kappa(h)}.$$

Then

$$W_f^{(n)} := \frac{1}{|G_n|} \sum_{h, \chi} \widehat{f}_n(h, \chi) U_h V_\chi \overline{\chi(h/2)},$$

with $(U_h \varphi)(g) := \varphi(g + h)$ and $(V_\chi \varphi)(g) := \chi(g)\varphi(g)$.

Let $f_n^{(2)}$ stand for the inverse Fourier transform of f_n with respect to the second variable; i.e.,

$$f_n^{(2)}(g, s) := \widehat{\Delta}_n \sum_{\kappa} f_n(g, \kappa)\kappa(s).$$

Then

$$(W_f^{(n)} \varphi)(s) = \Delta_n \sum_g f_n^{(2)}\left(\frac{s + g}{2}, g - s\right) \varphi(g).$$

7.7.11. *If $f \in \mathcal{S}_2(G \times \widehat{G})$ then the sequence $(W_f^{(n)})$ converges discretely to W_f with respect to the strong discrete approximant $((X_n, S_n))_{n \in \mathbb{N}}$ in 7.6.13 and this convergence is uniform.*

◁ The proof proceeds along the lines of 7.7.3. ▷

Propositions 7.7.4 and 7.7.5 are also applicable in this environment. Moreover, the proposition of 7.7.5 holds for the sequence $(W_f^{(n)})$, since each $W_f^{(n)}$ is hermitian as follows from the fact that f_n is a real function.

7.7.12. COMMENTS.

(1) The pseudodifferential operator A_f with symbol f is usually defined as

$$A_f := \int \int_{G \times \widehat{G}} \tilde{f}(h, \xi) V_h U_\xi d\widehat{\mu}(h) d\mu(\xi),$$

with $\tilde{f} := \mathcal{F}_G \otimes \mathcal{F}_{\widehat{G}}(f)$. It is easy to show by routine calculations with the formula $\int_{\widehat{G}} \chi(\xi) d\widehat{\mu}(\chi) = \delta(\xi)$, where $\int_G \varphi(\xi) \delta(\xi) d\mu(\xi) = \varphi(0)$, well-known in the distribution theory on locally compact abelian groups that if $\psi \in L_2(G)$ then the value $A_f \psi$ may be found by the formula in 7.7.1. It is also easy to justify these calculations rigorously, but this is immaterial for the aims of this section and so we use the definition of 7.7.1.

(2) Similar calculations lead to an analogous formula for the n th approximant:

$$A_f^{(n)} = \frac{1}{|G_n|} \sum_{g \in G_n, \chi \in \widehat{G}_n} \tilde{f}_n(\gamma\chi, g) X(g, \chi),$$

with $\tilde{f}_n := \mathcal{F}_{G_n} \otimes \mathcal{F}_{\widehat{G}_n}(f_n)$ and $f_n(g, \chi) := f(j_n(g), \widehat{j}_n(\chi))$.

(3) Note that in case $G := \mathbb{R}$ the symbol of (1) is not the Weyl symbol (symmetric symbol) of an operator but rather the so-called qp -symbol. See [506] about the qp -symbols of operators in $L_2(\mathbf{Q}_p^n)$ spaces.

The interrelation between the qp -symbols of A and A^* is not simple. Also, the conditions for the symbol f to make A_f selfadjoint are rather complicated. Moreover, these conditions do not guarantee that $A_f^{(n)}$ or $B_f^{(n)}$ will be selfadjoint for a selfadjoint A_f . The theory of pseudodifferential operators in $L_2(\mathbb{R}^n)$ also considers symmetric or Weyl symbols. The operator W_f with Weyl symbol f is selfadjoint if and only if f is a real function [33].

(4) The content of this section is taken from the article [5] by Albeverio, Gordon, and Khrennikov.

Chapter 8

Exercises and Unsolved Problems

This chapter collects not only simple questions for drill but also topics for serious research intended mostly at the graduate and postgraduate levels. Some problems need a creative thought to clarify and specify them. In short, this selection is rather haphazard, appearing *in statu nascendi*.

The problems in [271, 273–276, 278, 280] are the core of this chapter.

8.1. Nonstandard Hulls and Loeb Measures

8.1.1. The concept of nonstandard hull, stemming from the seminal works by Luxemburg, is a topical object of intensive study.

Many interesting facts are now in stock on the structure of the nonstandard hulls of Banach spaces and topological vector spaces (cf. [71, 187, 188, 466]). However, much is still unravelled in the interaction of the main constructions and concepts of Banach space theory and the various instances of nonstandard hulls. There is no detailed description for the nonstandard hulls of many function and operator spaces we deal with in functional analysis. We will give a few relevant statements.

By $X^\#$ we denote the nonstandard hull of a normed space X ; i.e. the quotient space of the external subspace of limited elements $\text{ltd}(X)$ by the monad $\mu(X)$ of the neighborhood filter about the origin of X , cf. 6.1.1. The prerequisites to functional analysis may be found in [83, 84, 227, 397].

Problem 1. Find conditions for $X^\#$ to possess the Kreĭn–Milman property in terms of X .

Look at [274] for a close bunch of problems related to the Kreĭn–Milman Theorem and its abstraction to Kantorovich spaces.

Problem 2. Find conditions for $X^\#$ to possess the Radon–Nikodĕm property in terms of X .

Problem 3. Study other geometric properties of the nonstandard hull of a Banach space such as smoothness, rotundity, the Asplund property, etc.

Problem 4. What is the stalkwise nonstandard hull of a continuous (measurable) Banach bundle? The same problem for the corresponding space of continuous (measurable) sections.

Problem 5. Describe the nonstandard hulls of various classes of bounded linear operators such as Radon–Nikodým operators, radonifying operators, order summing, p -absolutely summing and similar operators, etc.

8.1.2. The vector space $M(\nu)$ of cosets of measurable functions on a finite measure space $(\Omega, \mathcal{B}, \nu)$ possesses the metric

$$\rho(f, g) := \int_{\Omega} \frac{|f - g|}{1 + |f - g|} d\nu.$$

Furnished with the metric topology, $M(\nu)$ becomes a topological vector space.

Consider the nonstandard hull $M(\nu)^{\#} := \text{ltd}(M(\nu))/\mu_{\rho}(0)$, with $\mu_{\rho}(0) := \{f \in M(\nu) : \rho(f, 0) \approx 0\}$ and $\text{ltd}(M(\nu)) := \{f \in M(\nu) : \varepsilon f \in \mu_{\rho}(0) \text{ for } \varepsilon \approx 0\}$. Let $(\Omega, \mathcal{B}_L, \nu_L)$ stand for the corresponding Loeb measure space. Then $M(\nu)^{\#}$ and $M(\nu_L)$ are isometric spaces.

Problem 6. What is the matter with the space of Bochner measurable vector-functions $M(\nu, X)$? The same problem for Gelfand measurable and Pettis measurable functions.

Assume that E is an order ideal in $M(\nu)$; i.e., E is a subspace of $M(\nu)$ and, given $f \in M(\nu)$ and $g \in E$, the inequality $|f| \leq |g|$ implies that $f \in E$. Denote by $E(X)$ the space of $f \in M(\nu, X)$ such that the function $v(f) : t \mapsto \|f(t)\|$ ($t \in Q$) belongs to E , implying identification of equivalent functions. If E is a Banach lattice then $E(X)$ is a Banach space under the mixed norm $\|f\| = \|v(f)\|_E$.

Problem 7. Describe the nonstandard hull of $E(X)$.

8.1.3. Assume that (X, Σ, μ) is a finite measure space. Consider a hyperfinite set $M \subset {}^*X$, satisfying $\mu(A) = |A \cap M|/|M|$. Let (M, S_L, ν_L) stand for the corresponding Loeb measure space.

Problem 8. Is it true that under a suitable embedding $\varphi : \Sigma/\mu \rightarrow S_L/\nu_L$ the regular subalgebra $\varphi(\Sigma/\mu)$ is splittable? If this is so, describe the internal sets that correspond to the complementary factor (presenting, so to say, the “purely nonstandard” members of S_L/ν_L).

Problem 9. The same problem for embedding an interval with Lebesgue measure in some Loeb measure space.

Problem 10. *The same problem for the spaces implied in Problems 70 and 71.*

8.1.4. Let $(X, \mathcal{A}, \lambda)$ and (Y, \mathcal{B}, ν) be standard finite measure spaces. A function $\mu : \mathcal{A} \times Y \rightarrow \mathbb{R}$ is a *random measure* if

- (1) the function $\mu(A, \cdot)$ is \mathcal{B} -measurable for all $A \in \mathcal{A}$;
- (2) the function $\mu(\cdot, y)$ is a finite positive measure on \mathcal{A} for ν -almost all $y \in Y$.

Problem 11. *Suggest a definition of random Loeb measure μ_L so that it serves as a random measure for $(X, \mathcal{A}_L, \lambda_L)$ and $(Y, \mathcal{B}_L, \nu_L)$.*

Problem 12. *Describe interplay between the integral operators $\int f(x) d\mu(x, \cdot)$ and $\int f(x) d\mu_L(x, \cdot)$? What is the analog of S -integrability here?*

Some solution to Problems 11 and 12, belonging to Troitskiĭ [490], is presented in Section 6.6.

Problem 13. *Suggest a definition of vector-lattice-valued Loeb measure (in the absence of any topology). Do it so that the random Loeb measure of Problem 11 correlates with the concept of Loeb measure for the vector measure $A \mapsto \mu(A, \cdot)$.*

8.1.5. The next three problems are invoked by the article [18], belonging to the theory of spaces of differentiable functions (cf. [78, 131, 132, 356, 357]).

Problem 14. *Suggest a nonlinear potential theory by using the concept of Loeb measure.*

Problem 15. *Suggest a version of nonstandard capacity theory.*

Problem 16. *Define and study the spaces of differential forms by using Loeb measure (cf. [131, 132]).*

8.2. Hyperapproximation and Spectral Theory

8.2.1. In Chapter 7 we saw that each locally compact abelian group admits hyperapproximation; moreover, this approximation agrees with the Pontryagin–van Kampen duality and the Fourier transform for the original group is approximated the discrete Fourier transform on an approximant. This explains interest in studying the case of noncommutative groups. We come to a new class of “approximable” locally compact groups. This class seemingly includes amenable groups; however, no precise description is known for this class of groups.

Problem 17. *Given a locally compact (not necessarily abelian) group G , construct hyperapproximants to bounded endomorphisms of the $L_2(G)$ space.*

8.2.2. Approximations for locally compact abelian groups allow us to construct hyperapproximants to pseudodifferential operators in the Hilbert space of square integrable functions on a locally compact abelian group. This was done for the Schrödinger-type operators and Hilbert–Schmidt operators in the case of a special class of groups with compact and open subgroups in [503]. Another approach was pursued in [405, 406, 527].

The latter is more general since it is confined to the spaces of functions on a locally compact group. However, the former leads to more refined results. Therefore, interplay between these two approaches seems promising. The intriguing problem arises of abstracting available results to other pseudodifferential operators on a locally compact group and constructing analogous approximants to operators in function spaces over other approximable groups.

Another bunch of problems consists in studying the limit behavior of spectra and eigenvalues of hyperapproximants to a pseudodifferential operator on a locally compact abelian group.

Problem 18. *Study the limit behavior of the spectrum and eigenvalues of a hyperapproximant to a Schrödinger-type operator with positive potential growing at infinity. The same problem for a Hilbert–Schmidt operator.*

Problem 19. *The same problem as in Problem 18 for a Schrödinger operator with periodic potential.*

Problem 20. *Study interplay between hyperapproximants to a locally compact abelian group and its Bohr compactification.*

Problem 21. *Construct approximants to a Schrödinger-type operator with almost periodic potential on using Problem 20 and study their convergence.*

Problem 22. *Study the limit behavior of the spectra of approximants in a boundary value problem for the Schrödinger operator in a rectangular domain of a finite-dimensional space.*

8.2.3. We now list a bunch of problems relating to approximants to operators in function spaces over a noncommutative locally compact group and convergence of these approximants.

Problem 23. *Approximate various irreducible representations of the Heisenberg group by using representations of approximating finite groups.*

Problem 24. *Given a Hilbert function space on the Heisenberg group, find approximants to the operators in the algebra spanned over multiplications by the matrix elements of irreducible representations and shifts.*

Problem 25. *The same as in Problem 24 for other approximable nilpotent groups and suitable matrix groups over local fields.*

Problem 26. Study the approximation problem for simple Lie groups.

Problem 27. Study methods for summation of divergent series over an approximable discrete group, basing on hyperapproximation of this group.

Problem 28. Study interplay between nonstandard summation methods of divergent series with nonstandard extensions of a densely defined operator.

8.2.4. Hyperapproximation of operators is not always determined from hyperapproximation of a locally compact group. Moreover, if the domain of the operator under study is a function space over a domain other than a group the above-presented scheme of hyperapproximation is not applicable in general. However, we may try to construct hyperapproximants on using the specifics of the domain of an operator. We list a few relevant problems. Observe that Problems 30 and 31 are formulated jointly with Pliev.

Problem 29. Suggest a theory of Fredholm determinants that rests on appropriate hyperapproximation.

Problem 30. Prove the Lidskiĭ Theorem of coincidence of the matrix and spectral traces of a trace-class operators by hyperapproximation.

Problem 31. Use nonstandard discretization methods for studying the spectral properties of operator pencils. In particular, find an analog of the Keldysh Theorem on completeness of the derived chains of operator pencils (cf. [236]).

Problem 32. Construct a hyperfinite-rank analog of the Radon transform [161] in the spirit of [140, 142, 144, 146] (see Chapter 7).

Problem 33. Apply hyperapproximation of the Radon transform to analyzing the discrete scanning schemes of computer tomography [375].

8.3. Combining Nonstandard Methods

8.3.1. We have mentioned elsewhere that there are various ways of combining nonstandard methods: we may proceed with infinitesimal construction inside a Boolean valued universe or we may seek for Boolean valued interpretation in the framework of some theory of internal or external sets; cf. [271] and 4.8–4.11. However, serious difficulties arise and it is not always clear how to obviate them. At the same time, successive application of nonstandard methods leads often to a success as in [277, 280, 282, 294].

Problem 34. Develop a combined “scalarization-discretization” technique of unifying various combinations of nonstandard methods.

Problem 35. Suggest a Boolean valued version of the concept of Loeb measure and the relevant integration theory. Study the respective classes of operators. In particular, invent some Kantorovich-space-valued Loeb measure.

Problem 36. Give Boolean valued interpretations of available nonstandard hulls. Study the corresponding “descended” nonstandard hulls.

Problem 37. Using various nonstandard methods, derive a combined transfer principle from finite-dimensional normed algebras to relevant classes of Banach algebras.

Problem 38. Using a combined technique of “scalarization-discretization,” construct some hyperapproximants to representations of locally compact groups.

8.3.2. Substituting the laws of intuitionistic logic for the logical part of ZF (cf. [121, 150, 479]), we come to intuitionistic set theory ZF_I . We may construct models for ZF_I by using a similar scheme. Namely, considering a complete Heyting lattice, study numeric systems inside Heyting valued models and the corresponding algebraic structures; cf. [120, 133, 201].

Problem 40. Study classical Banach spaces inside Heyting valued models; cf. [50].

Problem 41. Does some interpretation of Hilbert space theory inside Heyting valued models lead to a meaningful theory of Hilbert modules?

8.3.3. Consider the following claim.

Let X and Y be normed spaces. Assume given X_0 a subspace of X and T_0 a bounded linear operator from X_0 to Y . Then, to each $0 < \varepsilon \in \mathbb{R}$, there is a bounded linear extension T of T_0 to the whole of X such that $\|T\| \leq (1 + \varepsilon)\|T_0\|$.

The Hahn–Banach Theorem fails in constructive mathematics. However (cf. [37]), it is well known that the above claim holds for functionals; i.e., in the case of $Y = \mathbb{R}$. Consequently, this claim is valid for functionals inside every Heyting valued model.

The same claim holds in the classical sense, i.e., in the von Neumann universe for compact operators ranging in the space $C(Q)$ of continuous functions on a compact space Q (see [318]).

Problem 42. Does the affinity of the two extension theorems for a functional and a compact operator ensue from some transfer principle for Heyting valued models?

Problem 43. For which objects and problems of functional analysis and operator theory is there an effective transfer principle resting on the technique of Heyting valued models? Topoi? Sheaves? (Cf. [122] and the entire collection [180].)

8.3.4. Let B be a quantum logic (cf. [281]). If we define the functions $[\cdot \in \cdot]$ and $[\cdot = \cdot]$ by the formulas of [154, 2.1.4] and introduce the same truth values as in [154, 2.1.7] then all Axioms ZF_2 – ZF_6 and AC become valid inside the universe $\mathbb{V}^{(B)}$. Therefore, we may practice set theory inside $\mathbb{V}^{(B)}$. In particular, the reals inside $\mathbb{V}^{(B)}$ correspond to observables in a mathematical model of a quantum mechanical system (cf. [476]).

In [476] there is shown that if B is a quantum logic [281] then $\mathbb{V}^{(B)}$ serves for a certain quantum set theory. Studying quantum theories as logical systems is a challenging topic as well as constructing quantum set theory and developing the corresponding quantum mathematics. However, this area of research still leaves much to be discovered. Adequate mathematical tools and signposts reveal themselves most likely in the theory of von Neumann algebras and various “noncommutative” branches stemming from it such as noncommutative probability theory, noncommutative integration, etc.

Problem 44. *Is there any reasonable version of the transfer principle from measure (integral) theory to noncommutative measure (integral) theory resting on the model $\mathbb{V}^{(B)}$ of quantum set theory?*

Problem 45. *Suggest a noncommutative theory for Loeb measure; i.e., apply the construction of Loeb measure to a measure on a quantum logic.*

Problem 46. *Suggest a theory of noncommutative vector (center-valued) integration on a von Neumann algebra (AW^* -algebra) and study the relevant spaces of measurable and integrable elements by Boolean valued realization.*

Problem 47. *What properties of the quantum complex numbers (i.e., the complex numbers inside $\mathbb{V}^{(B)}$ for a quantum logic B) correspond to meaningful properties of a von Neumann algebra (AW^* -algebra)?*

8.3.5. Let E be a vector lattice.

An operator T from E to an arbitrary vector space F is called *disjointly additive* if $T(x_1+x_2) = T(x_1)+T(x_2)$ for all $x_1, x_2 \in E$ such that $x_1 \perp x_2$ (i. e., $x_1 \wedge x_2 = 0$). We denote by $\mathcal{U}(E, F)$ the set of all disjointly additive order bounded operators from E to F . The members of $\mathcal{U}(E, F)$ are *abstract Urysohn operators* (cf. [355]).

Assume that F is a Kantorovich space. As demonstrated in [355], we make $\mathcal{U}(E, F)$ into a Kantorovich space by furnishing $\mathcal{U}(E, F)$ with the following order: $S \geq 0$ if and only if $S(x) \geq 0$ for all $x \in E$, with $S_1 \geq S_2$ implying that $S_1 - S_2 \geq 0$.

A disjointly additive operator in a Kantorovich space which commutes with each band projection we call an *abstract Nemytskiĭ operator*.

Problem 48. *Apply the “scalarization-discretization” method to nonlinear integral Urysohn operators as well as to their abstract analogs, i.e., bounded disjointly additive operators.*

Problem 49. Give a Boolean valued interpretation of disjointly additive functional and study the corresponding class of nonlinear operators.

Problem 50. Leaning on Problem 49, describe the band that is generated by a positive disjointly additive operator.

Problem 51. Suggest some Boolean valued realization for an abstract Ne-myskiĭ operator and find its functional representation.

8.3.6. The next problem resembles a species of convex analysis. However, it reflects the principal difficulty that stems from nonuniqueness of the standard part operation and related infinitesimal constructions inside a Boolean valued universe.

Problem 52. Considering a standard Kantorovich space, describe the subdifferential ∂p of the operator $p(e) := \inf^* \{f \in E : f \geq e\}$.

8.4. Convex Analysis and Extremal Problems

8.4.1. We start with problems on extreme points.

Problem 53. Study the points infinitely close to extreme points of a subdifferential.

Problem 54. Find the Boolean valued status of the o -extreme points of a subdifferential [279].

Problem 55. Describe the external equivalences that are kept invariant under the Young–Fenchel transform (cf. [279]).

8.4.2. Assume that (Q, Σ, μ) is a measure space, X is a Banach space, and E is a Banach lattice. Let Y stand for some space of measurable vector functions $u : Q \rightarrow X$, with identification of equivalent functions. Suppose that $f : Q \times X \rightarrow E^\bullet$ is a convex mapping in the second variable $x \in X$ for almost all $t \in Q$, with the composite $t \mapsto f(t, u(t))$ measurable for all $u \in Y$. We may then define some integral operator I_f over Y by the formula

$$I_f(u) := \int_Q f(t, u(t)) d\mu(t) \quad (u \in Y).$$

We agree that $I_f(u) := +\infty$ if the vector function $f(\cdot, u(\cdot))$ fails to be integrable. Clearly, $I_f : Y \rightarrow E^\bullet$ is a convex operator. Convex analysis pays much attention to operators of this sort.

In particular, the problems are topical of describing the subdifferential $\partial I_f(u_0)$ and the Young–Fenchel transform $(I_f)^*$ also called the conjugate of $(I_f)^*$. As

regards the general properties of convex operators, see [279, 315]; see [55, 100, 315] about integral convex functionals (in the case of $E = \mathbb{R}$).

Using the results of 6.3 and 6.4, we present the integral functional I_f as follows

$$I_f(u) = \circ \left(\Delta \sum_{k=1}^N f(t_k, u(t_k)) \right) \quad (u \in Y).$$

Problem 56. Study the convex integral functional I_f by means of the above representation. In particular, derive some formulas for calculating the subdifferential $\partial I_f(u_0)$.

Problem 57. Study convex and nonconvex integrands and corresponding integral functionals by infinitesimal discretization,

8.4.3. Various selection theorems are listed among powerful tools for studying functionals like I_f . We now state two available results precisely (cf. [55, 100, 315]).

Assume that Q is a topological (measurable) space, and X is a Banach space. A correspondence $\Gamma \subset Q \times X$ is called *lower semicontinuous (measurable)* provided that $\Gamma^{-1}(G)$ is open (measurable) for all open $G \subset X$. A mapping $\gamma : \text{dom}(f) \rightarrow X$ is a *selection* from Γ provided that $\gamma(q) \in \Gamma(q)$ for all $q \in \text{dom}(\Gamma)$.

Michael Continuous Selection Theorem. Suppose that Q is a paracompact space, Γ is lower semicontinuous correspondence, and $\Gamma(q)$ is a nonempty closed convex set for all $q \in Q$. Then there is a continuous selection from Γ .

Rokhlin–Kuratowski–Ryll–Nardzewski Theorem. Suppose that Q is a measurable space, X is a Polish space, i.e. a complete separable metric space, and $\Gamma \subset Q \times X$ is a measurable correspondence, with $\Gamma(q)$ closed for all $q \in Q$. Then there is a measurable selection from Γ .

Problem 58. Carry out hyperapproximation of a paracompact space and suggest a nonstandard proof of the Michael Continuous Selection Theorem.

Problem 59. Find a nonstandard approach to the measurable selection problem and, in particular, suggest a nonstandard proof for the Rokhlin–Kuratowski–Ryll–Nardzewski Theorem.

8.4.4. The following problems rest on the concept of infinitesimal optimum. As regards the prerequisites to convex analysis, see [100, 279, 422, 487].

Problem 60. Suggest a concept of infinitesimal solution to problems of optimal control and variational calculus.

Problem 61. Find an infinitesimal extension of an abstract nonlinear extremal problem with operator constraints and study the behavior of infinitesimal optima.

Problem 62. Pursue an infinitesimal approach to relaxation of nonconvex variational problems.

Problem 63. Suggest some subdifferential calculus for functions over Boolean algebras and study the extremal problems of optimal choice of some member of a Boolean algebra.

8.5. Miscellany

In the subsection we collect a few groups of problems related to various areas of mathematics.

8.5.1. RELATIVE STANDARDNESS.

Problem 64. Using the Euler broken lines with higher infinitesimal meshsize as compared with an infinitesimal ε in the van der Pol equation, find a direct proof of existence of “canards”—duck-shaped solutions—avoiding change-of-variable (passage to the Lenard plane) (cf. [542]).

Consider another definition of relative standardness:

$$x : st : y \iff (\exists^{st} f)(x = f(y)).$$

This definition implies that there is a natural $n : st : y$ succeeding some naturals nonstandard relative to y . This leads to a model of infinitesimal analysis with the “perforated” set of naturals which satisfies the transfer principle and the implication to the right in the idealization principle.

Problem 65. Suggest a reasonable axiomatics for such a version of infinitesimal analysis.

Assume that y is an admissible set and (X, Σ, μ) is a y -standard space with σ -additive measure μ . An element x in X is called y -random provided that $x \notin A$ for every y -standard set $A \in \Sigma$ satisfying $\mu(A) = 0$.

From 6.4.2(1) we infer the following

Theorem. If (X_1, Σ_1, μ_1) and (X_2, Σ_2, μ_2) are standard finite measure spaces, ξ_1 is a random element in X_1 , and ξ_2 is a ξ_1 -random element in X_2 ; then (ξ_1, ξ_2) is a random element in the product $X_1 \times X_2$.

Problem 66. Is the converse of the above theorem true?

Problem 67. Study properties of “dimensional” (“inhomogeneous”) real axis.

Problem 68. Is it possible to justify the physicists’ manipulations with fractional dimensions?

8.5.2. TOPOLOGY AND RADON MEASURES. Assume that X is an internal hyperfinite set and $\mathcal{R} \subset X^2$ is an equivalence on X which is the intersection of some family of k internal sets, with k a cardinal. Assume further that the nonstandard universe is k^+ -saturated (as usual, κ^+ is the least cardinal greater than κ). Furnish $X^\# := X/\mathcal{R}$ with the topology possessing $\{F^\# : F \subset X; F \text{ is internal}\}$ a base for the collection of closed sets. Then $X^\#$ is compact if and only if to each internal $A \supset \mathcal{R}$ there is a standardly-finite subset K of X of standardly finite size such that $X = A(K)$ where $A(K) := \{y \in X : (x, y) \in A \text{ for some } x \in K\}$. Moreover, every compact set may be presented in this manner.

Problem 69. Using these terms, describe connected, simply connected, disconnected, and extremally disconnected compact spaces.

Problem 70. Is each Radon measure on $X^\#$ induced by some Loeb measure on X ? In other words, is it true that to each Radon measure μ on $X^\#$ there is a Loeb measure ν_L on X such that $A \subset X^\#$ is μ -measurable if and only if $\pi^{-1}(A)$ is ν_L -measurable, with $\mu(A) = \nu_L(\pi^{-1}(A))$ (here $\pi : X \rightarrow X^\#$ stands for the quotient mapping).

It is well known that to each compact space \mathcal{X} there are an internal hyperfinite set X and an internal mapping $\Phi : X \rightarrow {}^*\mathcal{X}$ satisfying

$$(\forall^{st} \xi \in {}^*\mathcal{X})(\exists x \in X)(\Phi(x) \approx \xi);$$

moreover, if $\mathcal{R} := \{(x, y) : \Phi(x) \approx \Phi(y)\}$ then X/\mathcal{R} is homeomorphic with ${}^*\mathcal{X}$.

Problem 71. Is it true that to each Radon measure μ on \mathcal{X} there are some Φ satisfying the above conditions (or for all these Φ) and some Loeb measure ν_L on X (induced by an internal function $\nu : X \rightarrow {}^*\mathbb{R}$ “measuring the atoms”) such that

$$\int_{\mathcal{X}} f d\mu = \overset{\circ}{\left(\sum_{x \in X} {}^*f(\Phi(x))\nu(x) \right)}$$

for all bounded almost continuous functions f ?

Problem 72. Describe other topological properties of $X^\#$ (regularity, local compactness, etc.) in terms of the properties of \mathcal{R} . What other types of space may be obtained in the same manner?

Problem 73. Study the monads that serve as external preorders (i.e., quasi-uniform spaces).

8.5.3. THEORY OF ENTIRE FUNCTIONS. The next bunch of problems was suggested by Gordon.

Problem 74. Describe the class of nonstandard polynomials whose shadows are entire functions or entire functions of finite degree σ .

Recall that if f is an internal function from *X to ${}^*\mathbb{C}$ such that $|f({}^*x)|$ is limited for all $x \in X$ then the *shadow* or *standard part* of f is the function ${}^\circ f$ from X to \mathbb{C} such that $({}^\circ f)(x) = {}^\circ f(x)$.

Problem 75. Interpret the Paley–Wiener Theorem [428] in terms of Problem 74.

Problem 76. Find infinitesimal proofs for the Kotelnikoff Theorem and other interpolation theorems for entire functions [237, 314].

Problem 77. Using expansion of polynomials, derive the theorems on product expansion of entire functions (similar to the Eulerian expansion of $\sin x$) [217, 328].

8.5.4. ERGODIC THEORY. The bunch of Problems 78–83 is suggested by Kachurovskii [275].

Let N be an unlimited natural. A numeric sequence $\{x_n\}_{n=0}^N$ is called *microconvergent* if there is some real x^* such that $x_n \approx x^*$ for all unlimited $n \leq N$. Assume that a sequence $\{x_n\}_{n=0}^\infty$ converges in the conventional sense. The following three cases determine three types of convergence:

- (1) *White convergence*: the sequence $\{x_n\}_{n=0}^N$ microconverges for all unlimited N ;
- (2) *Color convergence*: there are two unlimited naturals N and M such that the sequence $\{x_n\}_{n=0}^N$ is microconvergent whereas the sequence $\{x_n\}_{n=0}^M$ is not;
- (3) *Black convergence*: the sequence $\{x_n\}_{n=0}^N$ is not microconvergent for every unlimited N .

Von Neumann Ergodic Theorem. Let U be an isometry of a complex Hilbert space H and let H_U be the fixed point subspace of U , i.e., $H_U = \{f \in H : Uf = f\}$. Denote the orthoprojection to H_U by P_U . Then

$$\lim_{n \rightarrow \infty} \left\| \frac{1}{n+1} \sum_{k=0}^n U^k f - P_U f \right\|_H = 0$$

for all $f \in H$.

Corollary. Assume that (Ω, λ) is a finite measure space, T is an automorphism of this space, and $f \in L_2(\Omega)$. Then the sequence $\left\{ \frac{1}{n+1} \sum_{k=0}^n f(T^k x) \right\}_{n=0}^\infty$ converges in the norm of $L_2(\Omega)$.

Let $\hat{L}_1(\Omega)$ stand for the external set of the elements $f \in L_1(\Omega)$ such that $\|f\|_1 \ll \infty$ and $\lambda(E) \approx 0$ implies $\int_E f d\lambda \approx 0$ for all $E \subset \Omega$. We also put $\hat{L}_2(\Omega) = \{f \in L_2(\Omega) : f^2 \in \hat{L}_1(\Omega)\}$. The next result is established in [205]; see [191, 192] for related topics.

Theorem of Bounded Fluctuation. If f belongs to $\hat{L}_2(\Omega)$ then the sequence of averages of f has bounded fluctuation (and consequently, its convergence is white or color, that is, nonblack).

Problem 78. Find other (possibly weaker) sufficient conditions that imply bounded fluctuation and nonblack convergence for the sequence of averages in the above corollary.

Problem 79. Find necessary conditions implying bounded fluctuation and nonblack convergence for the sequence in the above corollary which are as close as possible to the sufficient conditions of Problem 78.

Problem 80. The same as Problem 78 for the von Neumann Statistical Ergodic Theorem.

Problem 81. The same as Problem 79 for the von Neumann Statistical Ergodic Theorem and Problem 80.

Problem 82. The same as Problem 79 for the Birkhoff–Khinchin Ergodic Theorem.

Problem 83. The same as Problem 79 for the Birkhoff–Khinchin Ergodic Theorem and Problem 82.

8.5.5. We now list a few problems belonging to none of the above bunches.

Problem 84. Find criteria for nearstandardness and prenearstandardness for the elements of concrete classical normed spaces.

Problem 85. Develop the theory of bornological spaces resting on the monad of a bornology [182].

Problem 86. Find comparison tests for finite sums with infinitely many terms.

Problem 87. Construct approximation schemata for general algebraic (Boolean valued) systems.

Let X be a Banach space and let B be a complete Boolean algebra. Denote by $B[X]$ a completion inside $\mathbb{V}^{(B)}$ of the metric space X^\wedge , the standard name of X .

Problem 88. Find Banach spaces X and Boolean algebras B satisfying the equality $B[X'] = B[X]'$ inside $\mathbb{V}^{(B)}$.

Appendix

Here we sketch the theory of Boolean valued models of set theory in brief. More complete introductions are available in [29, 263, 276, 281].

A.1. Let B stand for a distinguished complete Boolean algebra. By a *Boolean valued interpretation* of an n -ary predicate P on a class X we mean any mapping $R : X^n \rightarrow B$.

Suppose that \mathcal{L} is a first-order language with some predicates P_0, P_1, \dots, P_n , and let R_0, R_1, \dots, R_n stand for some Boolean valued interpretations of these predicates on a class X .

Given a formula $\varphi(u_1, \dots, u_m)$ of the language \mathcal{L} and elements $x_1, \dots, x_m \in X$, we define the *truth value* $\llbracket \varphi(x_1, \dots, x_m) \rrbracket \in B$ by usual induction on the length of φ .

Dealing with atomic formulas, we put

$$\llbracket P_k(x_1, \dots, x_m) \rrbracket := R_k(x_1, \dots, x_m).$$

The steps of induction use the following rules:

$$\llbracket \varphi \vee \psi \rrbracket := \llbracket \varphi \rrbracket \vee \llbracket \psi \rrbracket,$$

$$\llbracket \varphi \wedge \psi \rrbracket := \llbracket \varphi \rrbracket \wedge \llbracket \psi \rrbracket,$$

$$\llbracket \varphi \rightarrow \psi \rrbracket := \llbracket \varphi \rrbracket \Rightarrow \llbracket \psi \rrbracket,$$

$$\llbracket \neg \varphi \rrbracket := \llbracket \varphi \rrbracket^*,$$

$$\llbracket (\forall x) \varphi \rrbracket := \bigwedge_{x \in X} \llbracket \varphi(x) \rrbracket,$$

$$\llbracket (\exists x) \varphi \rrbracket := \bigvee_{x \in X} \llbracket \varphi(x) \rrbracket,$$

with the symbols $\vee, \wedge, \Rightarrow, (\cdot)^*, \bigwedge, \bigvee$ on the right sides of the equalities designating the conventional Boolean operations on B and $a \Rightarrow b := a^* \vee b$.

A.2. A proposition $\varphi(x_1, \dots, x_m)$, with $x_1, \dots, x_m \in X$ and $\varphi(u_1, \dots, u_m)$ a formula, is *valid* (*true*, *veritable*, etc.) in an algebraic system $\mathbb{X} := (X, R_0, \dots, R_n)$ if $\llbracket \varphi(x_1, \dots, x_m) \rrbracket = \mathbf{1}$, where $\mathbf{1}$ is the greatest element of B . In this event we write $\mathbb{X} \models \varphi(x_1, \dots, x_m)$.

All logically true statements are valid in \mathbb{X} . If the predicate P_0 symbolizes equality then we require that the B -system $\mathbb{X} := (X, =, R_1, \dots, R_n)$ satisfies the axioms of equality. If this requirement is fulfilled then all logically true statements of the first-order logic with equality, expressible in the language $\mathcal{L} := \{=, P_1, \dots, P_n\}$, are valid in the B -system \mathbb{X} .

A.3. We now consider a Boolean valued interpretation on a class X of the language $\mathcal{L} := \{=, \in\}$ of ZFC, i.e., the first-order language \mathcal{L} with the two binary predicates: $=$ and \in . We denote the interpretations of these predicates by $\llbracket \cdot = \cdot \rrbracket$ and $\llbracket \cdot \in \cdot \rrbracket$, respectively. Thus, $\llbracket \cdot = \cdot \rrbracket, \llbracket \cdot \in \cdot \rrbracket : X \times X \rightarrow B$, and

$$\llbracket = (x, y) \rrbracket = \llbracket x = y \rrbracket, \quad \llbracket \in (x, y) \rrbracket = \llbracket x \in y \rrbracket \quad (x, y \in X).$$

Our nearest aim is to characterize the B -systems $\mathbb{X} := (X, \llbracket \cdot = \cdot \rrbracket, \llbracket \cdot \in \cdot \rrbracket)$ that model ZFC so that $\mathbb{X} \models \text{ZFC}$. The last condition amounts to the fact that all axioms of ZFC are valid in \mathbb{X} . So, for instance, by the rules of 1.2.1, the validity of the axiom of extensionality 1.1.4(1) means that, for all $x, y \in X$,

$$\llbracket x = y \rrbracket = \bigwedge_{z \in X} (\llbracket z \in x \rrbracket \Leftrightarrow \llbracket z \in y \rrbracket),$$

where $a \Leftrightarrow b := (a \Rightarrow b) \wedge (b \Rightarrow a)$ for all $a, b \in B$.

A.4. A B -system \mathbb{X} is called *separated* whenever for all $x, y \in X$ the statement $\llbracket x = y \rrbracket = \mathbf{1}$ implies $x = y$. An arbitrary B -system \mathbb{X} becomes separated after taking the quotient modulo the equivalence relation $\sim := \{(x, y) \in X^2 : \llbracket x = y \rrbracket = \mathbf{1}\}$. (This is done with the help of the well-known Frege–Russell–Scott trick; see [276].)

A B -system \mathbb{X} is said to be *isomorphic* to a B -system $\mathbb{X}' := (X', \llbracket \cdot = \cdot \rrbracket', \llbracket \cdot \in \cdot \rrbracket')$, if there is a bijection $\beta : X \rightarrow X'$ such that $\llbracket x = y \rrbracket = \llbracket \beta x = \beta y \rrbracket'$ and $\llbracket x \in y \rrbracket = \llbracket \beta x \in \beta y \rrbracket'$ for all $x, y \in X$.

A.5. Theorem. *There is a unique B -system \mathbb{X} up to isomorphism such that*

- (1) \mathbb{X} is separated;
- (2) The axioms of equality are valid in \mathbb{X} ;
- (3) The axiom of extensionality and the axiom of regularity hold in \mathbb{X} ;
- (4) If a function $f : \text{dom}(f) \rightarrow B$ satisfies $\text{dom}(f) \in \mathbb{V}$ and $\text{dom}(f) \subset \mathbb{X}$, then

$$\llbracket y \in x \rrbracket = \bigvee_{z \in \text{dom}(f)} (z) \wedge \llbracket z = y \rrbracket \quad (y \in \mathbb{X})$$

- for some $x \in \mathbb{X}$;
- (5) For each $x \in \mathbb{X}$, there is a function $f : \text{dom}(f) \rightarrow B$ with $\text{dom}(f) \in \mathbb{V}$, $\text{dom}(f) \subset \mathbb{X}$, such that equality holds in (4) for all $y \in \mathbb{X}$.

A.6. A B -system enjoying A.5 (1–5) is called a *Boolean valued model* of set theory and is denoted by the symbol $\mathbb{V}^{(B)} := (\mathbb{V}^{(B)}, \llbracket \cdot = \cdot \rrbracket, \llbracket \cdot \in \cdot \rrbracket)$. The class $\mathbb{V}^{(B)}$ is also called the *Boolean valued universe* over B . The basic properties of $\mathbb{V}^{(B)}$ are formulated as follows:

- (1) **TRANSFER PRINCIPLE.** Every axiom, and hence every theorem, of ZFC is valid in $\mathbb{V}^{(B)}$; in symbols, $\mathbb{V}^{(B)} \models \text{ZFC}$.
- (2) **MIXING PRINCIPLE.** If $(b_\xi)_{\xi \in \Xi}$ is a partition of unity in B , and $(x_\xi)_{\xi \in \Xi}$ is a family of elements of $\mathbb{V}^{(B)}$, then there is a unique element $x \in \mathbb{V}^{(B)}$ satisfying $b_\xi \leq \llbracket x = x_\xi \rrbracket$ for all $\xi \in \Xi$.

The element x is called the *mixing* of $(x_\xi)_{\xi \in \Xi}$ by $(b_\xi)_{\xi \in \Xi}$ and is denoted by $\text{mix}_{\xi \in \Xi} b_\xi x_\xi$.

- (3) **MAXIMUM PRINCIPLE.** For every formula $\varphi(u)$ of ZFC, possibly with constants from $\mathbb{V}^{(B)}$, there is an element $x_0 \in \mathbb{V}^{(B)}$ satisfying

$$\llbracket (\exists u)\varphi(u) \rrbracket = \llbracket \varphi(x_0) \rrbracket.$$

It follows in particular that if $\llbracket (\exists! x)\varphi(x) \rrbracket = 1$, then there is a unique x_0 in $\mathbb{V}^{(B)}$ satisfying $\llbracket \varphi(x_0) \rrbracket = 1$.

A.7. There is a unique mapping $x \mapsto x^\wedge$ from \mathbb{V} to $\mathbb{V}^{(B)}$ obeying the following conditions:

- (1) $x = y \leftrightarrow \llbracket x^\wedge = y^\wedge \rrbracket = 1$; $x \in y \leftrightarrow \llbracket x^\wedge \in y^\wedge \rrbracket = 1$ ($x, y \in \mathbb{V}$),
- (2) $\llbracket z \in y^\wedge \rrbracket = \bigvee_{x \in y} \llbracket x^\wedge = z \rrbracket$ ($z \in \mathbb{V}^{(B)}$, $y \in \mathbb{V}$).

This mapping is called the *canonical embedding* of \mathbb{V} into $\mathbb{V}^{(B)}$ and x^\wedge is referred to as the *standard name* of x .

- (3) **Restricted Transfer Principle.** Let $\varphi(u_1, \dots, u_n)$ be some restricted formula, i.e., the quantifiers of $\varphi(u_1, \dots, u_n)$ have the form $(\forall u)(u \in v \rightarrow \dots)$ or $(\exists u)(u \in v \wedge \dots)$ abbreviated to $(\forall u \in v)$ and $(\exists u \in v)$. Then

$$\varphi(x_1, \dots, x_n) \leftrightarrow \mathbb{V}^{(B)} \models \varphi(x_1^\wedge, \dots, x_n^\wedge)$$

for all $x_1, \dots, x_n \in \mathbb{V}$.

A.8. Given an element $X \in \mathbb{V}^{(B)}$, we define its *descent* $X \downarrow$ as $X \downarrow := \{x \in \mathbb{V}^{(B)} : \llbracket x \in X \rrbracket = 1\}$. The descent of X is a *cyclic* set; i.e., $X \downarrow$ is closed under mixing. More precisely, if $(b_\xi)_{\xi \in \Xi}$ is a partition of unity in B and $(x_\xi)_{\xi \in \Xi}$ is a family of elements of $X \downarrow$, then the mixing $\text{mix}_{\xi \in \Xi} b_\xi x_\xi$ lies in $X \downarrow$.

A.9. Let F be a correspondence from X to Y inside $\mathbb{V}^{(B)}$, i.e., $X, Y, F \in \mathbb{V}^{(B)}$ and $\llbracket F \subset X \times Y \rrbracket = \llbracket F \neq \emptyset \rrbracket = \mathbb{1}$. There is a unique correspondence $F\downarrow$ from $X\downarrow$ to $Y\downarrow$ satisfying $F(A)\downarrow = F\downarrow(A\downarrow)$ for every set $A \subset X\downarrow$ inside $\mathbb{V}^{(B)}$. Furthermore, $\llbracket F$ is a mapping from X to $Y \rrbracket = \mathbb{1}$ if and only if $F\downarrow$ is a mapping from $X\downarrow$ to $Y\downarrow$.

In particular, a function $f : Z^\wedge \rightarrow Y$ inside $\mathbb{V}^{(B)}$, where $Z \in \mathbb{V}$, defines its descent $f\downarrow : Z \rightarrow Y\downarrow$ by $f\downarrow(z) = f(z^\wedge)$ for all $z \in Z$.

A.10. We suppose that $X \in \mathcal{P}(\mathbb{V}^{(B)})$ and define a function $f : \text{dom}(f) \rightarrow B$ by putting $\text{dom}(f) = X$ and $\text{im}(f) = \{\mathbb{1}\}$. By A.5 (4) there is an element $X\uparrow \in \mathbb{V}^{(B)}$ satisfying

$$\llbracket y \in X\uparrow \rrbracket = \bigvee_{x \in X} \llbracket x = y \rrbracket \quad (y \in \mathbb{V}^{(B)}).$$

The element $X\uparrow$, unique by the axiom of extensionality, is called the *ascent* of X . Moreover, the following are true:

- (1) $Y\downarrow\uparrow = Y$ ($Y \in \mathbb{V}^{(B)}$),
- (2) $X\uparrow\downarrow = \text{mix}(X)$ ($X \in \mathcal{P}(\mathbb{V}^{(B)})$),

where $\text{mix}(X)$ consists of all mixings of the form $\text{mix } b_\xi x_\xi$, with $(x_\xi) \subset X$ and (b_ξ) a partition of unity in B .

A.11. Assume that $X, Y \in \mathcal{P}(\mathbb{V}^{(B)})$ and let F be a correspondence from X to Y . The following are equivalent:

- (1) There is a unique correspondence $F\uparrow$ from $X\uparrow$ to $Y\uparrow$ inside $\mathbb{V}^{(B)}$ such that $\text{dom}(F\uparrow) = \text{dom}(F)\uparrow$ and

$$F\uparrow(A\uparrow) = F(A)\uparrow$$

for every subset A of $\text{dom}(F)$;

- (2) The correspondence F is *extensional*, i.e.,

$$y_1 \in F(x_1) \rightarrow \llbracket x_1 = x_2 \rrbracket \leq \bigvee_{y_2 \in F(x_2)} \llbracket y_1 = y_2 \rrbracket.$$

A correspondence F is a mapping from X to Y if and only if $\llbracket F\uparrow : X\uparrow \rightarrow Y\uparrow \rrbracket = \mathbb{1}$. In particular, a mapping $f : Z \rightarrow Y\downarrow$ generates a function $f\uparrow : Z^\wedge \rightarrow Y$ such that $\llbracket f\uparrow(x^\wedge) = f(x) \rrbracket = \mathbb{1}$ for all $x \in Z$.

A.12. We assume that a nonempty set X carries some *B-structure*; i.e., we assume given a mapping $d : X \times X \rightarrow B$ satisfying the “metric axioms”:

- (1) $d(x, y) = 0 \leftrightarrow x = y$;
- (2) $d(x, y) = d(y, x)$;
- (3) $d(x, y) \leq d(x, z) \vee d(z, y)$.

Then there are an element $\mathcal{X} \in \mathbb{V}^{(B)}$ and an injection $\iota : X \rightarrow X' := \mathcal{X} \downarrow$ such that $d(x, y) = \llbracket \iota(x) \neq \iota(y) \rrbracket$ and every element $x' \in X'$ may be represented as $x' = \text{mix } b_\xi \iota x_\xi$, with $(x_\xi) \subset X$ and (b_ξ) a partition of unity in B . This fact enables us to consider sets with B -structure as subsets of $\mathbb{V}^{(B)}$ and to handle them according to the rules described above.

A.13. COMMENTS.

Boolean valued analysis (the term was coined by Takeuti) is a branch of functional analysis which uses Boolean valued models of set theory [473, 475]. Since recently this term has been treated in a broader sense implying the tools that rest on simultaneous use of two distinct Boolean valued models.

It was the Cohen method of forcing whose comprehension led to the invention of Boolean valued models of set theory which is attributed to the efforts by Scott, Solovay, and Vopěnka (see [29, 195, 276, 426, 511]).

A more detailed information on the topic of this Appendix can be found in [29, 148, 276, 281]. The machinery of Boolean valued models is framed as the technique of ascending and descending which suits the problems of analysis [263, 276, 281]. The embedding of the sets with Boolean structure into a Boolean valued universe leans on the Solovay–Tennenbaum method for complete Boolean algebras [453]

References

1. Akilov G. P. and Kutateladze S. S., Ordered Vector Spaces [in Russian], Nauka, Novosibirsk (1978).
2. Aksoy A. G. and Khamisi M. A., Nonstandard Methods in Fixed Point Theory, Springer-Verlag, Berlin etc. (1990).
3. Albeverio S., Fenstad J. E., et al., Nonstandard Methods in Stochastic Analysis and Mathematical Physics, Academic Press, Orlando etc. (1990).
4. Albeverio S., Luxemburg W. A. J., and Wolff M. P. H. (eds.), Advances in Analysis, Probability and Mathematical Physics: Contributions of Nonstandard Analysis, Kluwer Academic Publishers, Dordrecht (1995).
5. Albeverio S., Gordon E. I., and Khrennikov A. Yu., “Finite dimensional approximations of operators in the Hilbert spaces of functions on locally compact abelian groups,” *Acta Appl. Math.*, **64**, No. 1, 33–73 (2000).
6. Alekseev M. A., Glebskiĭ L. Yu., and Gordon E. I., On Approximations of Groups, Group Actions, and Hopf Algebras [Preprint, No. 491], Inst. Applied Fiz. Ros. Akad. Nauk, Nizhniĭ Novgorod (1999).
7. Alexandrov A. D., Problems of Science and the Researcher’s Standpoint [in Russian], Nauka (Leningrad Branch), Leningrad (1988).
8. Alexandrov A. D., “A general view of mathematics,” in: *Mathematics: Its Content, Methods and Meaning*. Ed. by A. D. Alexandrov, A. N. Kolmogorov, and M. A. Lavrent’ev. Dover Publications, Mineola and New York (1999). (Reprint of the 2nd 1969 ed.)
9. Aliprantis C. D. and Burkinshaw O., *Locally Solid Riesz Spaces*, Academic Press, New York etc. (1978).
10. Aliprantis C. D. and Burkinshaw O., *Positive Operators*, Academic Press, New York (1985).
11. Anderson R. M., “A nonstandard representation of Brownian motion and Itô integration,” *Israel J. Math. Soc.*, **25**, 15–46 (1976).
12. Anderson R. M., “Star-finite representations of measure spaces,” *Trans. Amer. Math. Soc.*, **271**, 667–687 (1982).

13. Andreev P. V., "On the standardization principle in the theory of bounded sets," *Vestnik Moskov. Univ. Ser. I Mat. Mekh.*, No. 1, 68–70 (1997).
14. Andreev P. V. and Gordon E. I., "The nonstandard theory of classes," *Vladikavkaz. Mat. Zh.*, **1**, No. 4, 2–16 (1999). (<http://alanianet.ru/omj/journal.htm>)
15. Anselone P. M., *Collectively Compact Operator Approximation. Theory and Applications to Integral Equations*, Englewood Cliffs, Prentice Hall (1971).
16. Archimedes, *Opera Omnia*, Teubner, Stuttgart (1972).
17. Arens R. F. and Kaplansky I., "Topological representation of algebras," *Trans. Amer. Math. Soc.*, **63**, No. 3, 457–481 (1948).
18. Arkeryd L. and Bergh J., "Some properties of Loeb–Sobolev spaces," *J. London Math. Soc.*, **34**, No. 2, 317–334 (1986).
19. Arkeryd L. O., Cutland N. J., and Henson W. (eds.), *Nonstandard Analysis. Theory and Applications (Proc. NATO Advanced Study Institute on Nonstandard Anal. and Its Appl., International Centre for Mathematical Studies, Edinburgh, Scotland, 30 June–13 July)*, Kluwer Academic Publishers, Dordrecht etc. (1997).
20. Arveson W., "Operator algebras and invariant subspaces," *Ann. of Math.*, **100**, No. 2, 433–532 (1974).
21. Arveson W., *An Invitation to C^* -Algebras*, Springer-Verlag, Berlin etc. (1976).
22. Attouch H., *Variational Convergence for Functions and Operators*, Pitman, Boston etc. (1984).
23. Aubin J.-P. and Ekeland I., *Applied Nonlinear Analysis*, Wiley-Interscience, New York (1979).
24. Aubin J.-P. and Frankowska H., *Set-Valued Analysis*, Birkhäuser-Verlag, Boston (1990).
25. Auslander L. and Tolimieri R., "Is computing with finite Fourier transform pure or applied mathematics?" *Bull. Amer. Math. Soc.*, **1**, No. 6, 847–897 (1979).
26. Bagarello F., "Nonstandard variational calculus with applications to classical mechanics. I: An existence criterion," *Internat. J. Theoret. Phys.*, **38**, No. 5, 1569–1592 (1999).
27. Bagarello F., "Nonstandard variational calculus with applications to classical mechanics. II: The inverse problem and more," *Internat. J. Theoret. Phys.*, **38**, No. 5, 1593–1615 (1999).
28. Ballard D., *Foundational Aspects of "Non"-Standard Mathematics*, Amer. Math. Soc., Providence, RI (1994).
29. Bell J. L., *Boolean-Valued Models and Independence Proofs in Set Theory*, Clarendon Press, New York etc. (1985).

30. Bell J. L., *A Primer of Infinitesimal Analysis*, Cambridge University Press, Cambridge (1998).
31. Bell J. L. and Slomson A. B., *Models and Ultraproducts: an Introduction*, North-Holland, Amsterdam etc. (1969).
32. Berberian S. K., *Baer *-Rings*, Springer-Verlag, Berlin (1972).
33. Berezin F. A. and Shubin M. A., *The Schrödinger Equation*, Kluwer Academic Publishers, Dordrecht (1991).
34. Berkeley G., *The Works*. Vol. 1–4, Thoemmes Press, Bristol (1994).
35. Bigard A., Keimel K., and Wolfenstein S., *Groupes et Anneaux Réticulés*, Springer-Verlag, Berlin etc. (1977). (Lecture Notes in Math., **608**.)
36. Birkhoff G., *Lattice Theory*, Amer. Math. Soc., Providence (1967).
37. Bishop E. and Bridges D., *Constructive Analysis*, Springer-Verlag, Berlin etc. (1985).
38. Blekhman I. I., Myshkis A. D., and Panovko A. G., *Mechanics and Applied Mathematics. Logics and Peculiarities of Mathematical Applications* [in Russian], Nauka, Moscow (1983).
39. Blumenthal L. M., *Theory and Applications of Distance Geometry*, Clarendon Press, Oxford (1953).
40. Bogolyubov A. N., “Read on Euler: he is a teacher for all of us,” *Nauka v SSSR*, No. 6, 98–104 (1984).
41. Boole G., *An Investigation of the Laws of Thought on Which Are Founded the Mathematical Theories of Logic and Probabilities*, Dover Publications, New York (1957).
42. Boole G., *Selected Manuscripts on Logic and Its Philosophy*, Birkhäuser-Verlag, Basel (1997). (Science Networks. Historical Studies, **20**.)
43. Borel E., *Probability et Certitude*, Presses Univ. de France, Paris (1956).
44. Bourbaki N., *Set Theory* [in French], Hermann, Paris (1958).
45. Boyer C. B., *A History of Mathematics*, John Wiley & Sons Inc., New York etc. (1968).
46. Bratteli O. and Robinson D., *Operator Algebras and Quantum Statistical Mechanics*, Springer-Verlag, New York etc. (1982).
47. Bukhvalov A. V., “Order bounded operators in vector lattices and in spaces of measurable functions,” *J. Soviet Math.*, **54**, No. 5, 1131–1176 (1991).
48. Bukhvalov A. V., Veksler A. I., and Geïler V. A., “Normed lattices,” *J. Soviet Math.*, **18**, 516–551 (1982).
49. Bukhvalov A. V., Veksler A. I., and Lozanovskii G. Ya., “Banach lattices: some Banach aspects of their theory,” *Russian Math. Surveys*, **34**, No.2, 159–212 (1979).
50. Burden C. W. and Mulvey C. J., “Banach spaces in categories of sheaves,” in: *Applications of Sheaves* (Proc. Res. Sympos. Appl. Sheaf Theory to

- Logic, Algebra and Anal., Durham, 1977), Springer-Verlag, Berlin etc., 1979, pp. 169–196.
51. Canjar R. M., “Complete Boolean ultraproducts,” *J. Symbolic Logic*, **52**, No. 2, 530–542 (1987).
 52. Cantor G., *Works on Set Theory* [in Russian], Nauka, Moscow (1985).
 53. Capinski M. and Cutland N. J., *Nonstandard Methods for Stochastic Fluid Mechanics*, World Scientific Publishers, Singapore etc. (1995). (Series on Advances in Mathematics for Applied Sciences, Vol. 27.)
 54. Carnot L., *Reflections on Metaphysics of Infinitesimal Calculus* [Russian translation], ONTI, Moscow (1933).
 55. Castaing C. and Valadier M., *Convex Analysis and Measurable Multifunctions*, Springer-Verlag, Berlin etc. (1977). (Lecture Notes in Math., **580**.)
 56. Chang C. C. and Keisler H. J., *Model Theory*, North-Holland, Amsterdam (1990).
 57. Chilin V. I., “Partially ordered involutive Baer algebras,” in: *Contemporary Problems of Mathematics. Newest Advances* [in Russian], VINITI, Moscow, **27**, 1985, pp. 99–128.
 58. Church A., *Introduction to Mathematical Logic*, Princeton University Press, Princeton (1956).
 59. Ciesielski K., *Set Theory for the Working Mathematician*, Cambridge University Press, Cambridge (1997).
 60. Clarke F. H., “Generalized gradients and applications,” *Trans. Amer. Math. Soc.*, **205**, No. 2, 247–262 (1975).
 61. Clarke F. H., *Optimization and Nonsmooth Analysis*, New York, Wiley (1983).
 62. Cohen P. J., *Set Theory and the Continuum Hypothesis*, Benjamin, New York etc. (1966).
 63. Cohen P. J., “On foundations of set theory,” *Uspekhi Mat. Nauk*, **29**, No. 5, 169–176 (1974).
 64. Courant R., *A Course of Differential and Integral Calculus*. Vol. 1 [Russian translation], Nauka, Moscow (1967).
 65. Courant R. and Robbins H., *What Is Mathematics? An Elementary Approach to Ideas and Methods*. Oxford University Press, Oxford etc. (1978).
 66. Cozart D. and Moore L. C. Jr., “The nonstandard hull of a normed Riesz space,” *Duke Math. J.*, **41**, 263–275 (1974).
 67. Cristiant C., “Der Beitrag Gödels für die Rechfertigung der Leibnizschen Idee von der Infinitesimalen,” *Österreich. Akad. Wiss. Math.-Natur. Kl. Sitzungsber. II*, Bd 192, No. 1–3, 25–43 (1983).
 68. Curtis T., *Nonstandard Methods in the Calculus of Variations*, Pitman, London (1993). (Pitman Research Notes in Mathematics, **297**.)
 69. Cutland N. J., “Nonstandard measure theory and its applications,” *Bull. London Math. Soc.*, **15**, No. 6, 530–589 (1983).

70. Cutland N. J., "Infinitesimal methods in control theory, deterministic and stochastic," *Acta Appl. Math.*, **5**, No. 2, 105–137 (1986).
71. Cutland N. J. (ed.), *Nonstandard Analysis and Its Applications*, Cambridge University Press, Cambridge (1988).
72. Dacunha-Castelle D. and Krivine J.-L., "Applications des ultraproducts a l'etude des espaces et des algebres de Banach," *Studia Math.*, **41**, 315–334 (1972).
73. Dales H. and Woodin W., *An Introduction to Independence for Analysts*, Cambridge University Press, Cambridge (1987).
74. Dauben J. W., Abraham Robinson. *The Creation of Nonstandard Analysis, a Personal and Mathematical Odyssey*, Princeton University Press, Princeton (1995).
75. Davis M., *Applied Nonstandard Analysis*, John Wiley & Sons, New York etc. (1977).
76. Day M., *Normed Linear Spaces*, Springer-Verlag, New York and Heidelberg (1973).
77. de Jonge E. and van Rooij A. C. M., *Introduction to Riesz Spaces*, Mathematisch Centrum, Amsterdam (1977).
78. Dellacherie C., *Capacities and Stochastic Processes* [in French], Springer-Verlag, Berlin etc. (1972).
79. Dem'yanov V. F. and Vasil'ev L. V., *Nondifferentiable Optimization* [in Russian], Nauka, Moscow (1981).
80. Dem'yanov V. F. and Rubinov A. M., *Constructive Nonsmooth Analysis*, Verlag Peter Lang, Frankfurt/Main (1995).
81. Diener F. and Diener M., "Les applications de l'analyse non standard," *Recherche*, **20**, No. 1, 68–83 (1989).
82. Diener F. and Diener M. (eds.), *Nonstandard Analysis in Practice*, Springer-Verlag, Berlin etc. (1995).
83. Diestel J., *Geometry of Banach Spaces: Selected Topics*, Springer-Verlag, Berlin etc. (1975).
84. Diestel J. and Uhl J. J., *Vector Measures*, Amer. Math. Soc., Providence, RI (1977). (Math. Surveys; **15**.)
85. Digernes T., Husstad E., and Varadarajan V., "Finite approximations of Weyl systems," *Math. Scand.*, **84**, 261–283 (1999).
86. Digernes T., Varadarajan V., and Varadhan S., "Finite approximations to quantum systems," *Rev. Math. Phys.*, **6**, No. 4, 621–648 (1994).
87. Dinculeanu N., *Vector Measures*, VEB Deutscher Verlag der Wissenschaften, Berlin (1966).
88. Dixmier J., *C^* -Algebras and Their Representations* [in French], Gauthier-Villars, Paris (1964).

89. Dixmier J., *C*-Algebras*, North-Holland, Amsterdam, New York, and Oxford (1977).
90. Dixmier J., *Les Algebres d'Operateurs dans l'Espace Hilbertien (Algebres de von Neumann)*, Gauthier-Villars, Paris (1996).
91. Dolecki S., "A general theory of necessary optimality conditions," *J. Math. Anal. Appl.*, **78**, No. 12, 267–308 (1980).
92. Dolecki S., "Tangency and differentiation: marginal functions," *Adv. in Appl. Math.*, **11**, 388–411 (1990).
93. Dragalin A. G., "An explicit Boolean-valued model for nonstandard arithmetic," *Publ. Math. Debrecen*, **42**, No. 3–4, 369–389 (1993).
94. Dunford N. and Schwartz J. T., *Linear Operators. Vol. 1: General Theory*, John Wiley & Sons Inc., New York (1988).
95. Dunford N. and Schwartz J. T., *Linear Operators. Vol. 2: Spectral Theory. Selfadjoint Operators in Hilbert Space*, John Wiley & Sons Inc., New York (1988).
96. Dunford N. and Schwartz J. T., *Linear Operators. Vol. 3: Spectral Operators*, John Wiley & Sons Inc., New York (1988).
97. Dunham W., *Euler: The Master of Us All*, The Mathematical Association of America, Washington (1999). (The Dolciani Mathematical Expositions, **22**.)
98. Eda K., "A Boolean power and a direct product of abelian groups," *Tsukuba J. Math.*, **6**, No. 2, 187–194 (1982).
99. Eda K., "On a Boolean power of a torsion free abelian group," *J. Algebra*, **82**, No. 1, 84–93 (1983).
100. Ekeland I. and Temam R., *Convex Analysis and Variational Problems*, North-Holland, Amsterdam (1976).
101. Eklof P., "Theory of ultraproducts for algebraists," in: *Handbook of Mathematical Logic*, North-Holland, Amsterdam, 1977.
102. Ellis D., "Geometry in abstract distance spaces," *Publ. Math. Debrecen*, **2**, 1–25 (1951).
103. Emel'yanov È. Yu., "Invariant homomorphisms of nonstandard enlargements of Boolean algebras and vector lattices," *Siberian Math. J.*, **38**, No. 2, 244–252 (1997).
104. Engeler E., *Metamathematics of Elementary Mathematics [in German]*, Springer-Verlag, Berlin etc. (1983).
105. Ershov Yu. L. and Palyutin E. A., *Mathematical Logic [in Russian]*, Nauka, Moscow (1987).
106. Esenin-Vol'pin A. S., "Analysis of potential implementability," in: *Logic Studies [in Russian]*, Izdat. Akad. Nauk SSSR, Moscow, 1959, pp. 218–262.
107. Espanol L., "Dimension of Boolean valued lattices and rings," *J. Pure Appl. Algebra*, No. 42, 223–236 (1986).

108. Euclid's "Elements." Books 7–10 [Russian translation], Gostekhizdat, Moscow and Leningrad (1949).
109. Euler L., Introduction to Analysis of the Infinite. Book I [Russian translation], ONTI, Moscow (1936); [English translation], Springer-Verlag, New York etc. (1988).
110. Euler L., Differential Calculus [Russian translation], Gostekhizdat, Leningrad (1949).
111. Euler L., Integral Calculus. Vol. 1 [Russian translation], Gostekhizdat, Moscow (1950).
112. Euler L., Opera Omnia. Series Prima: Opera Mathematica. Vol. 1–29, Birkhäuser-Verlag, Basel etc.
113. Euler L., Foundations of Differential Calculus, Springer-Verlag, New York (2000).
114. Faith C., Algebra: Rings, Modules, and Categories. Vol. 1, Springer-Verlag, Berlin etc. (1981).
115. Fakhoury H., "Représentations d'opérateurs à valeurs dans $L^1(X, \Sigma, \mu)$," Math. Ann., **240**, No. 3, 203–212 (1979).
116. Farkas E. and Szabo M., "On the plausibility of nonstandard proofs in analysis," Dialectica, **38**, No. 4, 297–310 (1974).
117. Fattorini H. O., The Cauchy Problem, Addison-Wesley (1983).
118. Foster A. L., "Generalized 'Boolean' theory of universal algebras. I. Subdirect sums and normal representation theorems," Math. Z., **58**, No. 3, 306–336 (1953).
119. Foster A. L., "Generalized 'Boolean' theory of universal algebras. II. Identities and subdirect sums of functionally complete algebras," Math. Z., **59**, No. 2, 191–199 (1953).
120. Fourman M. P., "The logic of toposes," in: Handbook of Mathematical Logic, North-Holland, Amsterdam, 1977.
121. Fourman M. P. and Scott D. S., "Sheaves and logic," in: Applications of Sheaves (Proc. Res. Sympos. Appl. Sheaf Theory to Logic, Algebra and Anal., Durham, 1977), Springer-Verlag, Berlin etc., 1979, pp. 302–401.
122. Fourman M. P., Mulvey C. J., and Scott D. S. (eds.), Applications of Sheaves (Proc. Res. Sympos. Appl. Sheaf Theory to Logic, Algebra and Anal., Durham, 1977), Springer-Verlag, Berlin etc., 1979.
123. Fraenkel A. and Bar-Hillel I., Foundations of Set Theory, North-Holland, Amsterdam (1958).
124. Fuchs L., Partially Ordered Algebraic Systems, Pergamon Press, Oxford (1963).
125. Fuller R. V., "Relations among continuous and various noncontinuous functions," Pacific J. Math., **25**, No. 3, 495–509 (1968).

126. Gandy R. O., "Limitations to mathematical knowledge," in: Logic Colloquium-80, North-Holland, New York and London, 1982, pp. 129–146.
127. Georgescu G. and Voiculescu I., "Eastern model theory for Boolean-valued theories," *Z. Math. Logik Grundlag. Math.*, No. 31, 79–88 (1985).
128. Glazman I. M. and Lyubich Yu. I., *Finite-Dimensional Linear Analysis*, M.I.T. Press, Cambridge (1974).
129. Gödel K., "What is Cantor's continuum problem?" *Amer. Math. Monthly*, **54**, No. 9, 515–525 (1947).
130. Gödel K., "Compatibility of the axiom of choice and the generalized continuum hypothesis with the axioms of set theory," *Uspekhi Mat. Nauk*, **8**, No. 1, 96–149 (1948).
131. Gol'dshtein V. M., Kuz'minov V. I., and Shvedov I. A., "The Künneth formula for L_p -cohomologies of warped products," *Sibirsk. Mat. Zh.*, **32**, No. 5, 29–42 (1991).
132. Gol'dshtein V. M., Kuz'minov V. I., and Shvedov I. A., "On approximation of exact and closed differential forms by compactly-supported forms," *Sibirsk. Mat. Zh.*, **33**, No. 2, 49–65 (1992).
133. Goldblatt R., *Topoi: Categorical Analysis of Logic*, North-Holland, Amsterdam etc. (1979).
134. Goldblatt R., *Lectures on the Hyperreals: An Introduction to Nonstandard Analysis*, Springer-Verlag, New York etc. (1998).
135. Goodearl K. R., *Von Neumann Regular Rings*, Pitman, London (1979).
136. Gordon E. I., "Real numbers in Boolean-valued models of set theory and K -spaces," *Dokl. Akad. Nauk SSSR*, **237**, No. 4, 773–775 (1977).
137. Gordon E. I., " K -spaces in Boolean-valued models of set theory," *Dokl. Akad. Nauk SSSR*, **258**, No. 4, 777–780 (1981).
138. Gordon E. I., "To the theorems of identity preservation in K -spaces," *Sibirsk. Mat. Zh.*, **23**, No. 3, 55–65 (1982).
139. Gordon E. I., *Rationally Complete Semiprime Commutative Rings in Boolean Valued Models of Set Theory*, Gor'kiĭ, VINITI, No. 3286-83 Dep (1983).
140. Gordon E. I., "Nonstandard finite-dimensional analogs of operators in $L_2(\mathbb{R}^n)$," *Sibirsk. Mat. Zh.*, **29**, No. 2, 45–59 (1988).
141. Gordon E. I., "Relatively standard elements in E. Nelson's internal set theory," *Sibirsk. Mat. Zh.*, **30**, No. 1, 89–95 (1989).
142. Gordon E. I., "Hyperfinite approximations of locally compact Abelian groups," *Dokl. Akad. Nauk SSSR*, **314**, No. 5, 1044–1047 (1990).
143. Gordon E. I., *Elements of Boolean Valued Analysis* [in Russian], Gor'kiĭ State University Press, Gor'kiĭ (1991).
144. Gordon E. I., "Nonstandard analysis and compact Abelian groups," *Sibirsk. Mat. Zh.*, **32**, No. 2, 26–40 (1991).

145. Gordon E. I., "On Loeb measures," *Izv. Vyssh. Uchebn. Zaved. Mat.*, No. 2, 25–33 (1991).
146. Gordon E. I., *Nonstandard Methods in Commutative Harmonic Analysis*, Amer. Math. Soc., Providence, RI (1997).
147. Gordon E. I. and Lyubetskii V. A., "Some applications of nonstandard analysis in the theory of Boolean valued measures," *Dokl. Akad. Nauk SSSR*, **256**, No. 5, 1037–1041 (1981).
148. Gordon E. I. and Morozov S. F., *Boolean Valued Models of Set Theory* [in Russian], Gor'kiĭ State University, Gor'kiĭ (1982).
149. Grätzer G., *General Lattice Theory*, Birkhäuser-Verlag, Basel (1978).
150. Grayson R. J., "Heyting-valued models for intuitionistic set theory," *Applications of Sheaves* (Proc. Res. Sympos. Appl. Sheaf Theory to Logic, Algebra and Anal., Durham, 1977), Springer-Verlag, Berlin etc., 1979, pp. 402–414.
151. Gurarii V. P., *Group Methods in Commutative Harmonic Analysis* [in Russian], VINITI, Moscow (1988).
152. Gutman A. E., "Banach fibering in lattice normed space theory," in: *Linear Operators Compatible with Order* [in Russian], Sobolev Institute Press, Novosibirsk, 1995, pp. 63–211.
153. Gutman A. E., "Locally one-dimensional K -spaces and σ -distributive Boolean algebras," *Siberian Adv. Math.*, **5**, No. 2, 99–121 (1995).
154. Gutman A. E., Emel'yanov È. Yu., Kusraev A. G., and Kutateladze S. S., *Nonstandard Analysis and Vector Lattices*, Kluwer Academic Publishers, Dordrecht (2000).
155. Hallet M., *Cantorian Set Theory and Limitation of Size*, Clarendon Press, Oxford (1984).
156. Halmos P. R., *Lectures on Boolean Algebras*, Van Nostrand, Toronto, New York, and London (1963).
157. Hanshe-Olsen H. and Störmer E., *Jordan Operator Algebras*, Pitman Publishers Inc., Boston etc. (1984).
158. Harnik V., "Infinitesimals from Leibniz to Robinson time—to bring them back to school," *Math. Intelligencer*, **8**, No. 2, 41–47 (1986).
159. Hatcher W., "Calculus is algebra," *Amer. Math. Monthly*, **89**, No. 6, 362–370 (1989).
160. Heinrich S., "Ultraproducts of L_1 -predual spaces," *Fund. Math.*, **113**, No. 3, 221–234 (1981).
161. Helgason S., *The Radon Transform*, Birkhäuser-Verlag, Boston (1999).
162. Henle J. M. and Kleinberg E. M., *Infinitesimal Calculus*, Alpine Press, Cambridge and London (1979).
163. Henson C. W., "On the nonstandard representation of measures," *Trans. Amer. Math. Soc.*, **172**, No. 2, 437–446 (1972).

164. Henson C. W., "When do two Banach spaces have isometrically isomorphic nonstandard hulls?" *Israel J. Math.*, **22**, 57–67 (1975).
165. Henson C. W., "Nonstandard hulls of Banach spaces," *Israel J. Math.*, **25**, 108–114 (1976).
166. Henson C. W., "Unbounded Loeb measures," *Proc. Amer. Math. Soc.*, **74**, No. 1, 143–150 (1979).
167. Henson C. W., "Infinitesimals in functional analysis," in: *Nonstandard Analysis and Its Applications*, Cambridge University Press, Cambridge etc., 1988, pp. 140–181.
168. Henson C. W. and Keisler H. J., "On the strength of nonstandard analysis," *J. Symbolic Logic*, **51**, No. 2, 377–386 (1986).
169. Henson C. W. and Moore L. C. Jr. "Nonstandard hulls of the classical Banach spaces," *Duke Math. J.*, **41**, No. 2, 277–284 (1974).
170. Henson C. W. and Moore L. C. Jr., "Nonstandard analysis and the theory of Banach spaces," in: *Nonstandard Analysis. Recent Developments*, Springer-Verlag, Berlin etc., 1983, pp. 27–112. (Lecture Notes in Math., **983**.)
171. Henson C. W., Kaufmann M., and Keisler H. J., "The strength of nonstandard methods in arithmetic," *J. Symbolic Logic*, **49**, No. 34, 1039–1057 (1984).
172. Hermann R., "Supernear functions," *Math. Japon.*, **31**, No. 2, 320 (1986).
173. Hernandez E. G., "Boolean-valued models of set theory with automorphisms," *Z. Math. Logik Grundlag. Math.*, **32**, No. 2, 117–130 (1986).
174. Hewitt E. and Ross K., *Abstract Harmonic Analysis*. Vol. 1 and 2, Springer-Verlag, Berlin etc. (1994).
175. Hilbert D. and Bernays P., *Foundations of Mathematics. Logical Calculi and the Formalization of Arithmetic* [in Russian], Nauka, Moscow (1979).
176. *Hilbert's Problems* [in Russian], Nauka, Moscow (1969).
177. Hiriart-Urruty J.-B., "Tangent cones, generalized gradients and mathematical programming in Banach spaces," *Math. Oper. Res.*, **4**, No. 1, 79–97 (1979).
178. Hobbes T., *Selected Works*. Vol. 1 [Russian translation], Mysl', Moscow (1965).
179. Hoehle U., "Almost everywhere convergence and Boolean-valued topologies," in: *Topology*, Proc. 5th Int. Meet., Lecce/Italy 1990, Suppl. Rend. Circ. Mat. Palermo, II. Ser. 29, 1992, pp. 215–227.
180. Hofmann K. H. and Keimel K., "Sheaf theoretical concepts in analysis: Bundles and sheaves of Banach spaces, Banach $C(X)$ -modules," in: *Applications of Sheaves*, Proc. Res. Sympos., Durham, 1977, Springer-Verlag, Berlin etc., 1979, pp. 415–441.
181. Hofstadter D. R., *Gödel, Escher, Bach: an Eternal Golden Braid*, Vintage Books, New York (1980).
182. Hogbe-Nlend H., *Theorie des Bornologie et Applications*, Springer-Verlag, Berlin etc. (1971).

183. Horiguchi H., "A definition of the category of Boolean-valued models," *Comment. Math. Univ. St. Paul.*, **30**, No. 2, 135–147 (1981).
184. Horiguchi H., "The category of Boolean-valued models and its applications," *Comment. Math. Univ. St. Paul.*, **34**, No. 1, 71–89 (1985).
185. Hrbáček K., "Axiomatic foundations for nonstandard analysis," *Fund. Math.*, **98**, No. 1, 1–24 (1978).
186. Hrbáček K., "Nonstandard set theory," *Amer. Math. Monthly*, **86**, No. 8, 659–677 (1979).
187. Hurd A. E. (ed.), *Nonstandard Analysis. Recent Developments*, Springer-Verlag, Berlin (1983).
188. Hurd A. E. and Loeb H., *An Introduction to Nonstandard Analysis*, Academic Press, Orlando etc. (1985).
189. Il'in V. A., Sadovnichii V. A., and Sendov Bl. Kh., *Mathematical Analysis [in Russian]*, Nauka, Moscow (1979).
190. Ionescu Tulcea A. and Ionescu Tulcea C., *Topics in the Theory of Lifting*, Springer-Verlag, Berlin etc. (1969).
191. Ivanov V. V., "Geometric properties of monotone functions and probabilities of random fluctuations," *Siberian Math. J.*, **37**, No. 1, 102–129 (1996).
192. Ivanov V. V., "Oscillations of means in the ergodic theorem," *Dokl. Akad. Nauk*, **347**, No. 6, 736–738 (1996).
193. Jaffe A., "Ordering the universe: the role of mathematics," *SIAM Rev.*, **26**, No. 4, 473–500 (1984).
194. Jarník V., *Bolzano and the Foundations of Mathematical Analysis*, Society of Czechosl. Math. Phys., Prague (1981).
195. Jech T. J., *Lectures in Set Theory with Particular Emphasis on the Method of Forcing*, Springer-Verlag, Berlin (1971).
196. Jech T. J., *The Axiom of Choice*, North-Holland, Amsterdam etc. (1973).
197. Jech T. J., "Abstract theory of abelian operator algebras: an application of forcing," *Trans. Amer. Math. Soc.*, **289**, No. 1, 133–162 (1985).
198. Jech T. J., "First order theory of complete Stonean algebras (Boolean-valued real and complex numbers)," *Canad. Math. Bull.*, **30**, No. 4, 385–392 (1987).
199. Jech T. J., "Boolean-linear spaces," *Adv. in Math.*, **81**, No. 2, 117–197 (1990).
200. Jech T. J., *Set Theory*, Springer-Verlag, Berlin (1997).
201. Johnstone P. T., *Topos Theory*, Academic Press, London etc. (1977).
202. Johnstone P. T., *Stone Spaces*, Cambridge University Press, Cambridge and New York (1982).
203. Jordan P., von Neumann J., and Wigner E., "On an algebraic generalization of the quantum mechanic formalism," *Ann. Math.*, **35**, 29–64 (1944).
204. Kachurovskii A. G., "Boundedness of mean fluctuations in the statistic ergodic theorem," *Optimization*, No. 48(65), 71–77 (1990).

205. Kachurovskii A. G., "The rate of convergence in ergodic theorems," *Uspekhi Mat. Nauk*, **51**, No. 4, 73–124 (1996).
206. Kadison R. V. and Ringrose J. R., *Fundamentals of the Theory of Operator Algebras*, Vol. 1 and 2, Amer. Math. Soc., Providence, RI (1997). Vol. 3 and 4, Birkhäuser-Verlag, Boston (1991–1992).
207. Kalina M., "On the ergodic theorem within nonstandard models," *Tatra Mt. Math. Publ.*, **10**, 87–93 (1997).
208. Kalton N. J., "The endomorphisms of L_p ($0 \leq p \leq 1$)," *Indiana Univ. Math. J.*, **27**, No. 3, 353–381 (1978).
209. Kalton N. J., "Linear operators on L_p for $0 < p < 1$," *Trans. Amer. Math. Soc.*, **259**, No. 2, 319–355 (1980).
210. Kalton N. J., "Representation of operators between function spaces," *Indiana Univ. Math. J.*, **33**, No. 5, 639–665 (1984).
211. Kalton N. J., "Endomorphisms of symmetric function spaces," *Indiana Univ. Math. J.*, **34**, No. 2, 225–247 (1985).
212. Kamo S., "Nonstandard natural number systems and nonstandard models," *J. Symbolic Logic*, **46**, No. 2, 365–376 (1987).
213. Kanovei V. and Reeken M., "Internal approach to external sets and universes," *Studia Logica*, part 1: **55**, 227–235 (1995); part II: **55**, 347–376 (1995); part III: **56**, 293–322 (1996).
214. Kanovei V. and Reeken M., "Mathematics in a nonstandard world. I," *Math. Japon.*, **45**, No. 2, 369–408 (1997).
215. Kanovei V. and Reeken M., "A nonstandard proof of the Jordan curve theorem," *Real Anal. Exchange*, **24**, No. 1, 161–169 (1998).
216. Kanovei V. and Reeken M., "Extending standard models of ZFC to models of nonstandard set theories," *Studia Logica*, **64**, No. 1, 37–59 (2000).
217. Kanovei V. G., "On well-posedness of the Euler method of decomposition of sinus to infinite product," *Uspekhi Mat. Nauk*, **43**, No. 4, 57–81 (1988).
218. Kanovei V. G., "Nonsoluble hypothesis in the theory of Nelson's internal sets," *Uspekhi Mat. Nauk*, **46**, No. 6, 3–50 (1991).
219. Kanovei V. G., "A nonstandard set theory in the ϵ -language," *Mat. Zametki*, **70**, No. 1, 46–50 (2001).
220. Kant I., *Selected Works*. Vol. 3 [Russian translation], *Mysl'*, Moscow (1964).
221. Kantorovich L. V., "On semi-ordered linear spaces and their applications to the theory of linear operations," *Dokl. Akad. Nauk SSSR*, **4**, No. 1–2, 11–14 (1935).
222. Kantorovich L. V., "General forms of some classes of linear operations," *Dokl. Akad. Nauk SSSR*, **3**, No. 9, 101–106 (1936).
223. Kantorovich L. V., "On one class of functional equations," *Dokl. Akad. Nauk SSSR*, **4**, No. 5, 211–216 (1936).

224. Kantorovich L. V., "On some classes of linear operations," *Dokl. Akad. Nauk SSSR*, **3**, No. 1, 9–13 (1936).
225. Kantorovich L. V., "To the general theory of operations in semi-ordered spaces," *Dokl. Akad. Nauk SSSR*, **1**, No. 7, 271–274 (1936).
226. Kantorovich L. V., "On functional equations," *Transactions of Leningrad State University*, **3**, No. 7, 17–33 (1937).
227. Kantorovich L. V. and Akilov G. P., *Functional Analysis*, Pergamon Press, Oxford and New York (1982).
228. Kantorovich L. V., Vulikh B. Z., and Pinsker A. G., *Functional Analysis in Semiordeed Spaces [in Russian]*, Gostekhizdat, Moscow and Leningrad (1950).
229. Kaplansky I., "Projections in Banach algebras," *Ann. of Math. (2)*, **53**, 235–249 (1951).
230. Kaplansky I., "Algebras of type I," *Ann. of Math. (2)*, **56**, 460–472 (1952).
231. Kaplansky I., "Modules over operator algebras," *Amer. J. Math.*, **75**, No. 4, 839–858 (1953).
232. Kawai T., "Axiom systems of nonstandard set theory," *Logic Symposia, Proc. Conf. Hakone 1979/1980*, Springer-Verlag, Berlin etc., 1981, pp. 57–65.
233. Kawebuta Sh., "A 'non-standard' approach to the field equations in the functional form," *Osaka J. Math.*, **23**, No. 1, 55–67 (1986).
234. Keisler H. J., *Elementary Calculus: An Approach Using Infinitesimals*, Prindle, Weber, and Schmidt, Boston, Mass. (1976).
235. Keisler H. J., "An infinitesimal approach to stochastic analysis," *Mem. Amer. Math. Soc.*, **48**, 184 p. (1984).
236. Keldysh M. V., "On completeness of eigenfunctions of some classes of non-self-adjoint operators," *Uspekhi Mat. Nauk*, **26**, No. 4, 15–41 (1971).
237. Khurgin Ya. I. and Yakovlev V. P., *Compactly Supported Functions in Physics and Technology [in Russian]*, Nauka, Moscow (1971).
238. Kleene S., *Mathematical Logic*, John Wiley & Sons, New York etc. (1967).
239. Kline M., *Mathematical Thought From Ancient to Modern Times*, Oxford University Press, Oxford (1972).
240. Kolesnikov E. V., Kusraev A. G., and Malyugin S. A., *On Dominated Operators [Preprint, No. 26]*, Sobolev Institute of Mathematics, Novosibirsk (1988).
241. Kollatz L., *Functional Analysis and Numerical Mathematics [in German]*, Springer-Verlag, Berlin etc. (1964).
242. Kolmogorov A. N. and Dragalin A. G., *Mathematical Logic. Supplementary Chapters [in Russian]*, Moscow University Press, Moscow (1984).
243. Kopperman R., *Model Theory and Its Applications*, Allyn and Bacon, Boston (1972).
244. Kopytov V. M., *Lattice-Ordered Groups [in Russian]*, Nauka, Moscow (1984).

245. Kornfeld I. P., Sinai Ya. G., and Fomin S. V., *Ergodic Theory*, Springer-Verlag, Berlin (1982).
246. Korol' A. I. and Chilin V. I., "Measurable operators in a Boolean valued model of set theory," *Dokl. Akad. Nauk UzSSR*, No. 3, 7–9 (1989).
247. Kramosil I., "Comparing alternative definitions of Boolean-valued fuzzy sets," *Kybernetika*, **28**, No. 6, 425–443 (1992).
248. Krasnosel'skiĭ M. A., *Positive Solutions of Operator Equations*, P. Noordhoff Ltd., Groningen (1964).
249. Kreisel G., "Observations of popular discussions on foundations," in: *Axiomatic Set Theory. Proc. Symposia in Pure Math.*, Amer. Math. Soc., Providence, RI, 1971, **1**, pp. 183–190.
250. Krupa A., "On various generalizations of the notion of an \mathcal{F} -power to the case of unbounded operators," *Bull. Polish Acad. Sci. Math.*, **38**, 159–166 (1990).
251. Kudryk T. S. and Lyantse V. È., "Operator-valued charges on finite sets," *Mat. Stud.*, **7**, No. 2, 145–156 (1997).
252. Kudryk T. S., Lyantse V. È., and Chujko G. I., "Nearstandardness of finite sets," *Mat. Stud.*, **2**, 25–34 (1993).
253. Kudryk T. S., Lyantse V. È., and Chujko G. I., "Nearstandard operators," *Mat. Stud.*, **3**, 29–40 (1994).
254. Kuratowski K. and Mostowski A., *Set Theory*, North-Holland, Amsterdam etc. (1967).
255. Kuribayashi Y., "Fourier transform using nonstandard analysis," *RIMS Kokyuroku*, **975**, 132–144 (1996).
256. Kusraev A. G., "Subdifferentials of nonsmooth operators and necessary conditions for an extremum," *Optimization (Novosibirsk)*, No. 24, 75–117 (1980).
257. Kusraev A. G., "On a general method of subdifferentiation," *Dokl. Akad. Nauk SSSR*, **257**, No. 4, 822–826 (1981).
258. Kusraev A. G., "Boolean valued analysis of duality between universally complete modules," *Dokl. Akad. Nauk SSSR*, **267**, No. 5, 1049–1052 (1982).
259. Kusraev A. G., "General disintegration formulas," *Dokl. Akad. Nauk SSSR*, **265**, No. 6, 1312–1316 (1982).
260. Kusraev A. G., "On the subdifferentials of composites of sets and functions," *Sibirsk. Mat. Zh.*, **23**, No. 2, 116–127 (1982).
261. Kusraev A. G., "On Boolean valued convex analysis," *Mathematische Optimierung. Theorie und Anwendungen*. Wartburg/Eisenach, 1983, pp. 106–109.
262. Kusraev A. G., "On the subdifferential of a sum," *Sibirsk. Mat. Zh.*, **25**, No. 4, 107–110 (1984).
263. Kusraev A. G., *Vector Duality and Its Applications [in Russian]*, Nauka, Novosibirsk (1985).

264. Kusraev A. G., "Linear operators in lattice-normed spaces," in: Studies on Geometry in the Large and Mathematical Analysis. Vol. 9 [in Russian], Trudy Inst. Mat. (Novosibirsk), Novosibirsk, 1987, pp. 84–123.
265. Kusraev A. G., "Boolean valued analysis and JB -algebras," Siberian Math. J., **35**, No. 1, 114–122 (1994).
266. Kusraev A. G., "Dominated operators," in: Linear Operators Compatible with Order [in Russian], Sobolev Institute Press, Novosibirsk, 1995, pp. 212–292.
267. Kusraev A. G., Boolean Valued Analysis of Duality Between Involutive Banach Algebras [in Russian], North Ossetian University Press, Vladikavkaz (1996).
268. Kusraev A. G., Dominated Operators, Kluwer Academic Publishers, Dordrecht (2000).
269. Kusraev A. G. and Kutateladze S. S., "Analysis of subdifferentials by means of Boolean valued models," Dokl. Akad. Nauk SSSR, **265**, No. 5, 1061–1064 (1982).
270. Kusraev A. G. and Kutateladze S. S., Notes on Boolean Valued Analysis [in Russian], Novosibirsk University Press, Novosibirsk (1984).
271. Kusraev A. G. and Kutateladze S. S., "On combination of nonstandard methods. I," Sibirsk. Mat. Zh., **31**, No. 5, 111–119 (1990).
272. Kusraev A. G. and Kutateladze S. S., "Nonstandard methods for Kantorovich spaces," Siberian Adv. Math., **2**, No. 2, 114–152 (1992).
273. Kusraev A. G. and Kutateladze S. S., "Nonstandard methods in geometric functional analysis," Amer. Math. Soc. Transl. Ser. 2, **151**, 91–105 (1992).
274. Kusraev A. G. and Kutateladze S. S., "The Kreĭn–Milman theorem and Kantorovich spaces," Optimization (Novosibirsk), No. 51 (68), 5–18 (1992).
275. Kusraev A. G. and Kutateladze S. S., 55 Unsolved Problems of Nonstandard Analysis [in Russian], Novosibirsk University Press, Novosibirsk (1993).
276. Kusraev A. G. and Kutateladze S. S., Nonstandard Methods of Analysis, Kluwer Academic Publishers, Dordrecht etc. (1994).
277. Kusraev A. G. and Kutateladze S. S., "Boolean-valued introduction to the theory of vector lattices," Amer. Math. Soc. Transl. Ser. 2, **163**, 103–126 (1995).
278. Kusraev A. G. and Kutateladze S. S., "Nonstandard methods in functional analysis," in: Interaction Between Functional Analysis, Harmonic Analysis, and Probability Theory, Marcel Dekker Inc., New York, 1995, pp. 301–306.
279. Kusraev A. G. and Kutateladze S. S., Subdifferentials: Theory and Applications, Kluwer Academic Publishers, Dordrecht etc. (1995).
280. Kusraev A. G. and Kutateladze S. S., "On combined nonstandard methods in the theory of positive operators," Mat. Stud., **7**, No. 1, 33–40 (1997).

281. Kusraev A. G. and Kutateladze S. S., *Boolean Valued Analysis*, Kluwer Academic Publishers, Dordrecht etc. (1999).
282. Kusraev A. G. and Kutateladze S. S., "On combined nonstandard methods in functional analysis," *Vladikavkaz. Mat. Zh.*, **2**, No. 1 (2000). (<http://alania-net.ru/omj/journal.htm>)
283. Kutateladze S. S., "Infinitesimal tangent cones," *Sibirsk. Mat. Zh.*, **27**, No. 6, 67–76 (1985).
284. Kutateladze S. S., "Microlimits, microsums, and Toeplitz matrices," *Optimization (Novosibirsk)*, No. 35, 16–23 (1985).
285. Kutateladze S. S., "Nonstandard analysis of tangential cones," *Dokl. Akad. Nauk SSSR*, **284**, No. 3, 525–527 (1985).
286. Kutateladze S. S., "A variant of nonstandard convex programming," *Sibirsk. Mat. Zh.*, **27**, No. 4, 84–92 (1986).
287. Kutateladze S. S., "Cyclic monads and their applications," *Sibirsk. Mat. Zh.*, **27**, No. 1, 100–110 (1986).
288. Kutateladze S. S., *Fundamentals of Nonstandard Mathematical Analysis*. Vol. 1: Naive Foundations of Infinitesimal Methods; Vol. 2: Set-Theoretic Foundations of Nonstandard Analysis; Vol. 3: Monads in General Topology [in Russian] (1986).
289. Kutateladze S. S., "On nonstandard methods in subdifferential calculus," in: *Partial Differential Equations* [in Russian], Nauka, Novosibirsk, 1986, pp. 116–120.
290. Kutateladze S. S., "Epiderivatives determined by a set of infinitesimals," *Sibirsk. Mat. Zh.*, **28**, No. 4, 140–144 (1987).
291. Kutateladze S. S., "Infinitesimals and a calculus of tangents," in: *Studies in Geometry in the Large and Mathematical Analysis* [in Russian], Nauka, Novosibirsk, 1987, pp. 123–135.
292. Kutateladze S. S., "On topological notions pertaining to continuity," *Sibirsk. Mat. Zh.*, **28**, No. 1, 143–147 (1987).
293. Kutateladze S. S., "Monads of proultrafilters and extensional filters," *Sibirsk. Mat. Zh.*, **30**, No. 1, 129–133 (1989).
294. Kutateladze S. S., "On fragments of positive operators," *Sibirsk. Mat. Zh.*, **30**, No. 5, 111–119 (1989).
295. Kutateladze S. S., "The stances of nonstandard analysis," in: *Contemporary Problems of Analysis and Geometry* [in Russian], Nauka, Novosibirsk, 1989, pp. 153–182. (*Trudy Inst. Mat.*, **14**.)
296. Kutateladze S. S., "Credenda of nonstandard analysis," *Siberian Adv. Math.*, **1**, No. 1, 109–137 (1991).
297. Kutateladze S. S., "Nonstandard tools for convex analysis," *Math. Japon.*, **43**, No. 2, 391–410 (1996).

298. Kutateladze S. S. (ed.), *Vector Lattices and Integral Operators*, Kluwer Academic Publishers, Dordrecht (1996).
299. Kutateladze S. S., *Formalisms of Nonstandard Analysis* [in Russian], Novosibirsk University Press, Novosibirsk (1999).
300. Kutateladze S. S., *Fundamentals of Functional Analysis*, Kluwer Academic Publishers, Dordrecht (1996).
301. Kuz'mina I. S., "Lusitania and its creator," in: *Nauka v SSSR*, No. 1, 107–110 (1985).
302. Lacey H. E., *The Isometric Theory of Classical Banach Spaces*, Springer-Verlag, Berlin etc. (1974).
303. Lambek J., *Lectures on Rings and Modules*, Blaisdell, Waltham (1966).
304. Landers D. and Rogge L., *Nonstandard Analysis* [in German], Springer-Verlag, Berlin etc. (1994).
305. Lang S., *Algebra*, Addison-Wesley, Reading (1965).
306. Langwitz D., "Nicht-standard-mathematik, begründet durch eine Verallgemeinerung der Körpererweiterung," *Exposit. Math.*, **1**, 307–333 (1983).
307. Larsen R., *Banach Algebras, an Introduction*, Dekker, New York (1973).
308. Lavrent'ev M. A., "Nikolai Nikolaevich Luzin," *Uspekhi Mat. Nauk*, **29**, No. 5, 177–182 (1979).
309. Lavrent'ev M. A., *Science. Technical Progress. Personnel* [in Russian], Nauka, Novosibirsk (1980).
310. League J.-M., "Science, technology, and the world," *Nauka i Zhizn'*, No. 11, 3–11 (1986).
311. Leibniz G. W., "Nova Methodus pro Maximis et Minimis, Itemque Tangentibus, quae nec Fractals nec Irrationales Quattitates, et Singulare pro Illus Calculi Genus," *Uspekhi Mat. Nauk*, **3**, No. 1, 166–173 (1948).
312. Leibniz G. W., *Selected Works. Vol. 1* [Russian translation], Mysl', Moscow (1983).
313. Leibniz G. W., *Selected Works. Vol. 2* [Russian translation], Mysl', Moscow (1984).
314. Levin B. Ya., *Distribution of Roots of Entire Functions* [in Russian], Gostekhizdat, Moscow (1956).
315. Levin V. L., *Convex Analysis in Spaces of Measurable Functions and Its Application in Mathematics and Economics* [in Russian], Nauka, Moscow (1985).
316. Levy A., *Basic Set Theory*, Springer-Verlag, Berlin etc. (1979).
317. Li N., "The Boolean-valued model of the axiom system of GB," *Chinese Sci. Bull.*, **36**, No. 2, 99–102 (1991).
318. Lindenstrauss J., "Extension of compact operators," *Mem. Amer. Math. Soc.*, **48**, 112 p. (1964).

319. Lindenstrauss J. and Tzafriri L., Classical Banach Spaces. Vol. 1: Sequence Spaces, Springer-Verlag, Berlin etc. (1977).
320. Lindenstrauss J. and Tzafriri L., Classical Banach Spaces. Vol. 2: Function Spaces, Springer-Verlag, Berlin etc. (1979).
321. Locher J. L. (ed.), The World of M. C. Escher, Abradale Press, New York (1988).
322. Loeb P. A., "Conversion from nonstandard to standard measure spaces and applications to probability theory," *Trans. Amer. Math. Soc.*, **211**, 113–122 (1975).
323. Loeb P. A., "An introduction to nonstandard analysis and hyperfinite probability theory," in: *Probabilistic Analysis and Related Topics*. Vol. 2 (Ed. A. T. Bharucha-Reid), Academic Press, New York, 1979, pp. 105–142.
324. Loeb P. and Wolff M. P. H. (eds.), *Nonstandard Analysis for the Working Mathematician*, Kluwer Academic Publishers, Dordrecht etc. (2000).
325. Loewen P. D., *Optimal Control via Nonsmooth Analysis*, Amer. Math. Soc., Providence, RI (1993).
326. Lowen R., "Mathematics and fuzziness," in: *Fuzzy Sets Theory and Applications* (Louvain-la-Neuve, 1985), NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci., 177, Reidel, Dordrecht, and Boston, 1986, pp. 3–38.
327. Lutz R. and Gose M., *Nonstandard Analysis. A Practical Guide with Applications*, Springer-Verlag, Berlin etc. (1981). (Lecture Notes in Math., **881**.)
328. Luxemburg W. A. J., "A general theory of monads," in: *Applications of Model Theory to Algebra, Analysis and Probability*, Holt, Rinehart, and Winston, New York, 1969, pp. 18–86.
329. Luxemburg W. A. J., "A nonstandard approach to Fourier analysis," in: *Contributions to Nonstandard Analysis*, North-Holland, Amsterdam, 1972, pp. 16–39.
330. Luxemburg W. A. J. (ed.), *Applications of Model Theory to Algebra, Analysis and Probability*, Holt, Rinehart, and Winston, New York (1969).
331. Luxemburg W. A. J. and Robinson A. (eds.), *Contribution to Non-Standard Analysis*, North-Holland, Amsterdam (1972).
332. Luxemburg W. A. J. and Zaanen A. C., *Riesz Spaces*. Vol. 1, North-Holland, Amsterdam and London (1971).
333. Luzin N. N., "Real function theory: State of the art," in: *Proceedings of the All-Russia Congress of Mathematicians* (Moscow, April 27–May 4, 1927), Glavnauka, Moscow and Leningrad, 1928, pp. 11–32.
334. Luzin N. N., *The Theory of Real Functions* [in Russian], Uchpedgiz, Moscow (1940).
335. Luzin N. N., *Collected Works*. Vol. 3 [in Russian], Izdat. Akad. Nauk SSSR, Moscow (1959).

336. Luzin N. N., Differential Calculus [in Russian], Vysshaya Shkola, Moscow (1961).
337. "Luzin N. N. is an outstanding mathematician and teacher," Vestnik Akad. Nauk SSSR, No. 11, 95–102 (1984).
338. Lyantse V. È., "Is it possible to ignore nonstandard analysis?" in: General Theory of Boundary Problems [in Russian], Naukova Dumka, Kiev, 1983, pp. 108–112.
339. Lyantse V. È., "On nonstandard analysis," in: Mathematics of Today [in Russian], Vishcha Shkola, Kiev, 1986, pp. 26–44.
340. Lyantse V. È., "Nearstandardness on a finite set," Dissertationes Math., **369**, 63 p. (1997).
341. Lyantse V. È. and Kudrik T. S., "On functions of discrete variable," in: Mathematics of Today [in Russian], Vishcha Shkola, Kiev, 1987.
342. Lyantse W. and Kudryk T., Introduction to Nonstandard Analysis, VNTL Publishers, Lviv (1997).
343. Lyantse V. È. and Yavorskyj Yu. M., "Nonstandard Sturm–Liouville difference operator," Mat. Stud., **10**, No. 1, 54–68 (1998).
344. Lyantse V. È. and Yavorskyj Yu. M., "Nonstandard Sturm–Liouville difference operator. II," Mat. Stud., **11**, No. 1, 71–82 (1999).
345. Lyubetskiĭ V. A., "On some algebraic problems of nonstandard analysis," Dokl. Akad. Nauk SSSR, **280**, No. 1, 38–41 (1985).
346. Lyubetskiĭ V. A. and Gordon E. I., "Boolean completion of uniformities," in: Studies on Nonclassical Logics and Formal Systems [in Russian], Nauka, Moscow, 1983, pp. 82–153.
347. Lyubetskiĭ V. A. and Gordon E. I., "Immersing bundles into a Heyting valued universe," Dokl. Akad. Nauk SSSR, **268**, No. 4, 794–798 (1983).
348. MacLane S., Categories for the Working Mathematician, Springer-Verlag, New York (1971).
349. Mal'tsev A. I., Algebraic Systems. Posthumous Edition, Springer-Verlag, New York and Heidelberg (1973).
350. Malykhin V. I., "New aspects of general topology pertaining to forcing," Uspekhi Mat. Nauk, **43**, No. 4, 83–94 (1988).
351. Malykhin V. I. and Ponomarëv V. I., "General topology (set-theoretic direction)," in: Algebra. Topology. Geometry [in Russian], VINITI, Moscow, 1975, **13**, pp. 149–229.
352. Manin Yu. I., A Course in Mathematical Logic, Springer-Verlag, Berlin etc. (1977).
353. Marinakis K., "The infinitely small in space and in time," Trans. Hellenic Inst. Appl. Sci., No. 8, 62 p. (1970).
354. Martin-Löf P., "The definition of random sequences," Inform. and Control., **9**, 602–619 (1966).

355. Mazon J. M. and Segura de León S., "Order bounded orthogonally additive operators," *Rev. Roumaine Math. Pures Appl.*, **35**, No. 4, 329–354 (1990).
356. Maz'ya V. G., *The Spaces of S. L. Sobolev* [in Russian], Leningrad University Publ. House, Leningrad (1985).
357. Maz'ya V. G. and Khavin V. P., "A nonlinear potential theory," *Uspekhi Mat. Nauk*, **27**, No. 6, 67–138 (1972).
358. Mckee T., "Monadic characterizable in nonstandard topology," *Z. Math. Logik Grundlag. Math.*, **26**, No. 5, 395–397 (1940).
359. Medvedev F. A., "Cantor's set theory and theology," in: *Historio-Mathematical Studies* [in Russian], Nauka, Moscow, 1985, **29**, pp. 209–240.
360. Medvedev F. A., "Nonstandard analysis and the history of classical analysis," in: *Laws of the Development of Modern Mathematics* [in Russian], Nauka, Moscow, 1987, pp. 75–84.
361. Melter R., "Boolean valued rings and Boolean metric spaces," *Arch. Math.*, No. 15, 354–363 (1964).
362. Mendelson E., *Introduction to Mathematical Logic*, Van Nostrand, Princeton etc. (1963).
363. Milvay C. J., "Banach sheaves," *J. Pure Appl. Algebra*, **17**, No. 1, 69–84 (1980).
364. Mochover M. and Hirschfeld J., *Lectures on Non-Standard Analysis*, Springer-Verlag, Berlin etc. (1969). (Lecture Notes in Math., **94**.)
365. Molchanov I. S., "Set-valued estimators for mean bodies related to Boolean models," *Statistics*, **28**, No. 1, pp. 43–56 (1996).
366. Molchanov V. A., *The One-Dimensional Mathematical Analysis in Nonstandard Presentation* [in Russian], K. A. Fedin Saratov Ped. Inst., Saratov (1989).
367. Molchanov V. A., "The use of double nonstandard enlargements in topology," *Sibirsk. Mat. Zh.*, **30**, No. 3, 64–71 (1989).
368. Molchanov V. A., *Introduction into the Calculus of Infinitesimals* [in Russian], K. A. Fedin Saratov Ped. Inst., Saratov (1990).
369. Molchanov V. A., *Nonstandard Convergences in the Spaces of Mappings* [in Russian], K. A. Fedin Saratov Ped. Inst., Saratov (1991).
370. Monk J. D. and Bonnet R. (eds.), *Handbook of Boolean Algebras*. Vol. 1–3, North-Holland, Amsterdam etc. (1989).
371. Moore L. C. Jr., "Hyperfinite extensions of bounded operators on a separable Hilbert space," *Trans. Amer. Math. Soc.*, **218**, 285–295 (1976).
372. Mostowski A., *Constructible Sets with Applications*, North-Holland, Amsterdam (1969).
373. Murphy G., *C^* -Algebras and Operator Theory*, Academic Press, Boston, (1990).

374. Namba K., "Formal systems and Boolean valued combinatorics," in: South-east Asian Conference on Logic (Singapore, 1981), Stud. Logic Found. Math., 111, North-Holland, Amsterdam and New York, 1983, pp. 115–132.
375. Natterer F., The Mathematics of Computerized Tomography, Teubner and John Wiley & Sons, Stuttgart and Chichester, etc. (1986).
376. Naïmark M. A., Normed Rings, Noordhoff, Groningen (1959).
377. Naïmark M. A., Representation Theory of Groups [in Russian], Nauka, Moscow (1976).
378. Nelson E., "Internal set theory. A new approach to nonstandard analysis," Bull. Amer. Math. Soc., **83**, No. 6, 1165–1198 (1977).
379. Nelson E., Radically Elementary Probability Theory, Princeton University Press, Princeton (1987).
380. Nelson E., "The syntax of nonstandard analysis," Ann. Pure Appl. Logic, **38**, No. 2, 123–134 (1988).
381. Nepeïvoda N. N., Applied Logic [in Russian], Novosibirsk University Press, Novosibirsk (2000).
382. Neubrunn I., Riečan B., and Riečanou Z., "An elementary approach to some applications of nonstandard analysis," Rend. Circ. Math. Palermo, No. 3, 197–200 (1984).
383. Neumann J., von, Collected Works. Vol. 1 and 2, Pergamon Press, New York, Oxford, London, and Paris (1961).
384. Newton I., The Mathematical Papers of Isaac Newton [in Russian], ONTI, Moscow and Leningrad (1937).
385. Ng Siu-Ah, Hypermodels in Mathematical Finance. Modelling via Infinitesimal Analysis, World Scientific, Singapore etc. (2001).
386. Nishimura H., "Boolean valued and Stone algebra valued measure theories," Math. Logic Quart., **40**, No. 1, 69–75 (1994).
387. Novikov P. S., Constructive Mathematical Logic from the Point of View of Classical Logic [in Russian], Nauka, Moscow (1977).
388. Ozawa M., "Boolean valued analysis approach to the trace problem of AW^* -algebras," J. London Math. Soc. (2), **33**, No. 2, 347–354 (1986).
389. Ozawa M., "Embeddable AW^* -algebras and regular completions," J. London Math. Soc., **34**, No. 3, 511–523 (1986).
390. Ozawa M., "Boolean-valued interpretation of Banach space theory and module structures of von Neumann algebras," Nagoya Math. J., **117**, 1–36 (1990).
391. Pedersen G. K., Analysis Now, Springer-Verlag, Berlin etc. (1995).
392. Penot J.-P., "Compact nets, filters and relations," J. Math. Anal. Appl., **93**, No. 2, 400–417 (1983).
393. Péraire Y., "Une nouvelle théorie des infinitesimaux," C. R. Acad. Sci. Paris Ser. I., **301**, No. 5, 157–159 (1985).

394. Péraire Y., "A general theory of infinitesimals," *Sibirsk. Mat. Zh.*, **31**, No. 3, 107–124 (1990).
395. Péraire Y., "Théorie relative des ensembles internes," *Osaka J. Math.*, **29**, No. 2, 267–297 (1992).
396. Péraire Y., "Some extensions of the principles of idealization transfer and choice in the relative internal set theory," *Arch. Math. Logic.*, **34**, 269–277 (1995).
397. Pietsch A., *Operator Ideals*, VEB Deutschen Verlag der Wissenschaften, Berlin (1978).
398. Pinus A. G., *Boolean Constructions in Universal Algebras*, Kluwer Academic Publishers, Dordrecht etc. (1993).
399. Pirce R. S., "Modules over commutative regular rings," *Mem. Amer. Math. Soc.*, No. 70 (1967).
400. Pontryagin L. S., *Analysis of Infinitesimals* [in Russian], Nauka, Moscow (1980).
401. Pontryagin L. S., *Mathematical Analysis for Schoolchildren* [in Russian], Nauka, Moscow (1980).
402. Pontryagin L. S., *Continuous Groups* [in Russian], Nauka, Moscow (1984).
403. Postnikov A. G., *Introduction to Analytic Number Theory* [in Russian], Nauka, Moscow (1971).
404. Prokhorova M. F., "On relative nearstandardness in IST," *Siberian Math. J.*, **39**, No. 3, 518–521 (1998).
405. Raebiger F. and Wolff M. P. H., "On the approximation of positive operators and the behavior of the spectra of the approximants," in: *Integral Equations Operator Theory*, 1997, **28**, pp. 72–86.
406. Raebiger F. and Wolff M. P. H., "Spectral and asymptotic properties of dominated operators," *J. Austral. Math. Soc. Ser. A*, **63**, 16–31 (1997).
407. Rasiowa E. and Sikorski R., *The Mathematics of Metamathematics*, PWN, Warsaw (1962).
408. *Reader on the History of Mathematics* [in Russian], Prosveshchenie, Moscow (1977).
409. Reed M. and Simon B., *Methods of Modern Mathematical Physics*, Academic Press, New York and London (1972).
410. Reinhardt H.-J., *Analysis of Approximation Methods for Differential and Integral Equations*, Springer-Verlag, New York, Berlin, Heidelberg, and Tokyo (1985).
411. Rema P. S., "Boolean metrization and topological spaces," *Math. Japon.*, **9**, No. 9, 19–30 (1964).
412. Repicky M., "Cardinal characteristics of the real line and Boolean-valued models," *Comment. Math. Univ. Carolin.*, **33**, No. 1, 184 (1992).

413. Reshetnyak Yu. G., *A Course of Mathematical Analysis*. Vol. 1, Part 1 [in Russian], Sobolev Institute Press, Novosibirsk (1999).
414. Revuzhenko A. F., *Mechanics of Elastic-Plastic Media and Nonstandard Analysis* [in Russian], Novosibirsk University Press, Novosibirsk (2000).
415. Rickart Ch., *General Theory of Banach Algebras*, Van Nostrand, Princeton (1960).
416. Riesz F. and Szökefalvi-Nagy B., *Lectures on Functional Analysis* [in French], Académi Kiadó, Budapest (1972).
417. Robert A., *Analyse Non-Standard*, Queen's University Press, Kingston (1984).
418. Robert A., "One approche naive de l'analyse non-standard," *Dialectica*, **38**, 287–290 (1984).
419. Robinson A., *Introduction to the Theory of Models and to the Metamathematics of Algebra*, North-Holland, Amsterdam (1963).
420. Robinson A., "The metaphysics of the calculus," *Problems in the Philosophy of Mathematics*, North-Holland, Amsterdam, 1967, **1**, pp. 28–46.
421. Robinson A., *Non-Standard Analysis*, Princeton Univ. Press, Princeton (1996).
422. Rockafellar R. T., *Convex Analysis*, Princeton University Press, Princeton (1970).
423. Rockafellar R. T., *The Theory of Subgradients and Its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Springer-Verlag, Berlin etc. (1981).
424. Rockafellar R. T. and Wets R. J.-B., *Variational Analysis*, Springer-Verlag, Birkhäuser-Verlag (1998).
425. Rosser J. B., *Logic for Mathematicians*, McGraw-Hill voon Company Inc., New York etc. (1953).
426. Rosser J. B., *Simplified Independence Proofs. Boolean Valued Models of Set Theory*, Academic Press, New York and London (1969).
427. Rubio J. E., *Optimization and Nonstandard Analysis*, Marcel Dekker, New York and London (1994). (Pure and Applied Mathematics, **184**.)
428. Rudin W., *Functional Analysis*, McGraw-Hill Inc., New York (1991).
429. Ruzavin G. I., *Philosophical Problems of the Foundations of Mathematics* [in Russian], Nauka, Moscow (1983).
430. Sakai S. *C^* -Algebras and W^* -Algebras*, Springer-Verlag, Berlin etc. (1971).
431. Salbany S. and Todorov T., "Nonstandard analysis in topology: nonstandard and standard compactifications," *J. Symbolic Logic*, **65**, No. 4, 1836–1840 (2000).
432. Saracino D. and Weispfenning V., "On algebraic curves over commutative regular rings," in: *Model Theory and Algebra (a Memorial Tribute to Abraham Robinson)*, Springer-Verlag, New York etc., 1969. (Lecture Notes in Math., **498**.)

433. Sarymsakov T. A. et al., Ordered Algebras [in Russian], Fan, Tashkent (1983).
434. Schaefer H. H., Banach Lattices and Positive Operators, Springer-Verlag, Berlin etc. (1974).
435. Schröder J., "Das Iterationsverfahren bei allgemeinerem Abstands begriff," Math. Z., Bd 66, 111–116 (1956).
436. Schwartz L., Analysis. Vol. 1 [in French], Hermann, Paris (1967).
437. Schwarz H.-U., Banach Lattices and Operators, Teubner, Leipzig (1984).
438. Schwinger J., "Unitary operator bases. The special canonical group," Proc. Natl. Acad. Sci. USA, 46, 570–579, 1401–1405 (1960).
439. Severi F., "Italian algebraic geometry, its rigor, methods, and problems," Mathematics, 3, No. 1, 111–141 (1959).
440. Sextus Empiricus, Selected Works. Vol. 1 [Russian translation], Mysl', Moscow (1976).
441. Shamaev I. I., "On decomposition and representation of regular operators," Sibirsk. Mat. Zh., 30, No. 2, 192–202 (1989).
442. Shamaev I. I., "An operator disjoint from lattice homomorphisms, Maharam operators, and integral operators," Optimization (Novosibirsk), No. 46, 154–159 (1990).
443. Shamaev I. I., Topological Methods for the Theory of Regular Operators in Kantorovich Spaces: ScD Thesis [in Russian], Yakutsk (1994).
444. Shoenfield J. R., Mathematical Logic, Addison-Wesley, Reading (1967).
445. Shoenfield J. R., "Axioms of set theory," in: Handbook of Mathematical Logic, North-Holland, Amsterdam, 1977.
446. Shotaev G. N., "On bilinear operators in lattice-normed spaces," Optimization (Novosibirsk), No. 37, 38–50 (1986).
447. Sikorski R., Boolean Algebras, Springer-Verlag, Berlin etc. (1964).
448. Sikorskii M. R., "Some applications of Boolean-valued models to study operators on polynormed spaces," Sov. Math., 33, No. 2, 106–110 (1989).
449. Smith K., "Commutative regular rings and Boolean-valued fields," J. Symbolic Logic, 49, No. 1, 281–297 (1984).
450. Smithson R., "Subcontinuity for multifunctions," Pacific J. Math., 61, No. 4, 283–288 (1975).
451. Sobolev V. I., "On a poset valued measure of a set, measurable functions and some abstract integrals," Dokl. Akad. Nauk SSSR, 91, No. 1, 23–26 (1953).
452. Solovay R. M., "A model of set theory in which every set of reals is Lebesgue measurable," Ann. of Math. (2), 92, No. 2, 1–56 (1970).
453. Solovay R. and Tennenbaum S., "Iterated Cohen extensions and Souslin's problem," Ann. Math., 94, No. 2, 201–245 (1972).
454. Solov'ev Yu. P. and Troitskiĭ E. V., C^* -Algebras and Elliptic Operators in Differential Topology [in Russian], Factorial, Moscow (1996).

-
455. Sourour A. R., "Operators with absolutely bounded matrices," *Math. Z.*, **162**, No. 2, 183–187 (1978).
456. Sourour A. R., "The isometries of $L^p(\Omega, X)$," *J. Funct. Anal.*, **30**, No. 2, 276–285 (1978).
457. Sourour A. R., "A note on integral operators," *Acta Sci. Math.*, **41**, No. 43, 375–379 (1979).
458. Sourour A. R., "Pseudo-integral operators," *Trans. Amer. Math. Soc.*, **253**, 339–363 (1979).
459. Sourour A. R., "Characterization and order properties of pseudo-integral operators," *Pacific J. Math.*, **99**, No. 1, 145–158 (1982).
460. Stern J., "Some applications of model theory in Banach space theory," *Ann. Math. Logic*, **9**, No. 1, 49–121 (1976).
461. Stern J., "The problem of envelopes for Banach spaces," *Israel J. Math.*, **24**, No. 1, 1–15 (1976).
462. Stewart I., "Frog and Mouse revisited," *Math. Intelligencer*, **8**, No. 4, 78–82 (1986).
463. Stroyan K. D., "Infinitesimal analysis of curves and surfaces," in: *Handbook of Mathematical Logic*, North-Holland, Amsterdam, 1977.
464. Stroyan K. D., "Infinitesimal calculus in locally convex spaces. I," *Trans. Amer. Math. Soc.*, **240**, 363–384 (1978); "Locally convex infinitesimal calculus. II," *J. Funct. Anal.*, **53**, No. 1, 1–15 (1983).
465. Stroyan K. D. and Bayod J. M., *Foundations of Infinitesimal Stochastic Analysis*, North-Holland, Amsterdam etc. (1986).
466. Stroyan K. D. and Luxemburg W. A. J., *Introduction to the Theory of Infinitesimals*, Academic Press, New York etc. (1976).
467. Struik D. J., *A Concise History of Mathematics*, Dover Publications, New York (1987).
468. Stummel F., "Diskrete Konvergenz Linearer Operatoren. I," *Math. Ann.*, **190**, 45–92 (1970).
469. Stummel F., "Diskrete Konvergenz Linearer Operatoren. II," *Math. Z.*, **120**, 231–264 (1971).
470. Sunder V. S., *An Invitation to Von Neumann Algebras*, Springer-Verlag, New York etc. (1987).
471. Székely G. J., *Paradoxes in Probability Theory and Mathematical Statistics [in Russian]*, Mir, Moscow (1990).
472. Tacon D. G., "Nonstandard extensions of transformations between Banach spaces," *Trans. Amer. Math. Soc.*, **260**, No. 1, 147–158 (1980).
473. Takeuti G., *Two Applications of Logic to Mathematics*, Iwanami and Princeton University Press, Tokyo and Princeton (1978).
474. Takeuti G., "A transfer principle in harmonic analysis," *J. Symbolic Logic*, **44**, No. 3, 417–440 (1979).

475. Takeuti G., “Boolean valued analysis,” in: Applications of Sheaves (Proc. Res. Sympos. Appl. Sheaf Theory to Logic, Algebra and Anal., Durham, 1977), Springer-Verlag, Berlin etc., 1979, pp. 714–731.
476. Takeuti G., “Quantum set theory,” in: Current Issues in Quantum Logic (Erice, 1979), Plenum Press, New York and London, 1981, pp. 303–322.
477. Takeuti G., “ C^* -algebras and Boolean valued analysis,” Japan J. Math., **9**, No. 2, 207–246 (1983).
478. Takeuti G., “Von Neumann algebras and Boolean valued analysis,” J. Math. Soc. Japan, **35**, No. 1, 1–21 (1983).
479. Takeuti G. and Titani S., “Heyting-valued universes of intuitionistic set theory,” Logic Symposia, Proc. Conf. Hakone 1979/1980, Springer-Verlag, Berlin and New York, 1981, pp. 189–306. (Lecture Notes in Math., **89**.)
480. Takeuti G. and Titani S., “Globalization of intuitionistic set theory,” Ann. Pure Appl. Logic, **33**, No. 2, 195–211 (1987).
481. Takeuti G. and Zaring W. M., Introduction to Axiomatic Set Theory, Springer-Verlag, New York etc. (1971).
482. Takeuti G. and Zaring W. M., Axiomatic Set Theory, Springer-Verlag, New York (1973).
483. Tall D., “The calculus of Leibniz—an alternative modern approach,” Math. Intelligencer, **2**, No. 1, 54–60 (1979/1980).
484. Tanaka K. and Yamazaki T., “A nonstandard construction of Haar measure and weak Koenig’s lemma,” J. Symbolic Logic, **65**, No. 1, 173–186 (2000).
485. Thibault L., “Epidifferentielles de fonctions vectorielles,” C. R. Acad. Sci. Paris Sér. I Math., **290**, No. 2, A87–A90 (1980).
486. Thibault L., “Subdifferentials of nonconvex vector-valued mappings,” J. Math. Anal. Appl., **86**, No. 2 (1982).
487. Tikhomorov V. M., “Convex analysis,” in: Analysis. II. Convex Analysis and Approximation Theory, Encycl. Math. Sci. **14**, 1–92 (1990).
488. Topping D. M., “Jordan algebras of self-adjoint operators,” Mem. Amer. Math. Soc., **53**, 48 p. (1965).
489. Trenogin V. A., Functional Analysis [in Russian], Nauka, Moscow (1980).
490. Troitskii V. G., “Nonstandard discretization and the Loeb extension of a family of measures,” Siberian Math. J., **34**, No. 3, 566–573 (1993).
491. Tsalenko M. Sh. and Shul’geifer E. F., Fundamentals of Category Theory [in Russian], Nauka, Moscow (1974).
492. Uspenskii V. A., “Seven considerations on the philosophical topics of mathematics,” in: Laws of the Development of Modern Mathematics [in Russian], Nauka, Moscow, 1987, pp. 106–155.
493. Uspenskii V. A., What Is Nonstandard Analysis? [in Russian], Nauka, Moscow (1987).

494. Van de Berg I., *Nonstandard Asymptotic Analysis*, Springer-Verlag, Berlin etc. (1987).
495. van der Hoeven J., “On the computation of limsup,” *J. Pure Appl. Algebra*, **117–118**, 381–394 (1997).
496. Varadarajan V., *Quantization of Semisimple Groups and Some Applications* [Preprint] (1996).
497. Veksler A. I., “On a new construction of the Dedekind completion of some lattices and l -groups with division,” *Siberian Math. J.*, **10**, No. 6, 1206–1213 (1969).
498. Veksler A. I., “Banach cyclic spaces and Banach lattices,” *Dokl. Akad. Nauk SSSR*, **213**, No. 4, 770–773 (1973).
499. Veksler A. I. and Geĭler V. A., “On the order and disjoint completeness of semiordered linear spaces,” *Sibirsk. Mat. Zh.*, **13**, No. 1, 43–51 (1972).
500. Veksler A. I. and Gordon E. I., “Nonstandard extension of not-everywhere-defined positive operators,” *Sibirsk. Mat. Zh.*, **35**, No. 4, 640–647 (1994).
501. Venkataraman K., “Boolean valued almost periodic functions: existence of the mean,” *J. Indian Math. Soc.*, **43**, No. 1–4, 275–283 (1979).
502. Venkataraman K., “Boolean valued almost periodic functions on topological groups,” *J. Indian Math. Soc.*, **48**, No. 1–4, 153–164 (1984).
503. Vershik A. M. and Gordon E. I., “The groups locally embedded into the class of finite groups,” *Algebra and Analysis*, **9**, No. 1, 72–86 (1997).
504. Vilenkin N., “The commander of ‘Lusitania’,” *Znanie—Sila*, No. 1, 27–29 (1984).
505. Vladimirov D. A., *Boolean Algebras in Analysis*, Kluwer Academic Publishers, Dordrecht (2001).
506. Vladimirov V. S., Volovich I. V., and Zelenov E. I., *p -Adic Analysis and Mathematical Physics* [in Russian], Nauka, Moscow (1994).
507. Voltaire, *Verses and Proses* [Russian translation], Moskovskii Rabochii, Moscow (1997).
508. von Neumann J., *Collected Works. Vol. 1: Logic, Theory of Sets and Quantum Mechanics*, Pergamon Press, Oxford etc. (1961).
509. von Neumann J., *Collected Works. Vol. 3: Rings of Operators*, Pergamon Press, New York, Oxford, London, and Paris (1961).
510. von Neumann J., *Collected Works. Vol. 4: Continuous Geometry and Other Topics*, Pergamon Press, Oxford, London, New York, and Paris (1962).
511. Vopěnka P., “General theory of ∇ -models,” *Comment. Math. Univ. Carolin.*, **7**, No. 1, 147–170 (1967).
512. Vopěnka P., “The limits of sheaves over extremally disconnected compact Hausdorff spaces,” *Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys.*, **15**, No. 1, 1–4 (1967).

513. Vopěnka P., *Mathematics in the Alternative Set Theory*, Teubner, Leipzig (1979).
514. Vulikh B. Z., *Introduction to the Theory of Partially Ordered Spaces*, Noordhoff, Groningen (1967).
515. Vygodskii M. Ya., *Fundamentals of the Calculus of Infinitesimals* [in Russian], GTTI, Moscow and Leningrad (1933).
516. Wang Hao and McNaughton R., *Axiomatic Systems of Set Theory* [in French], Gauthier-Villars, Paris; E. Nauwelaerts, Louvain (1953).
517. Ward D. E., "Convex subcones of the contingent cone in nonsmooth calculus and optimization," *Trans. Amer. Math. Soc.*, **302**, No. 2, 661–682 (1987).
518. Ward D. E., "The quantification tangent cones," *Canad. J. Math.*, **40**, No. 3, 666–694 (1988).
519. Ward D. E., "Corrigendum to 'Convex subcones of the contingent cone in nonsmooth calculus and optimization,'" *Trans. Amer. Math. Soc.*, **311**, No. 1, 429–431 (1989).
520. Wattenberg F., "Nonstandard measure theory: Avoiding pathological sets," *Trans. Amer. Math. Soc.*, **250**, 357–368 (1979).
521. Weil A., "Euler," *Amer. Math. Monthly*, **91**, No. 9, 537–542 (1984).
522. Weis L., "Decompositions of positive operators and some of their applications," in: *Functional Analysis: Survey and Recent Results. Vol. 3: Proc. 3rd Conf., Paderborn, 24–29 May, 1983, Amsterdam, 1984*.
523. Weis L., "On the representation of order continuous operators by random measures," *Trans. Amer. Math. Soc.*, **285**, No. 2, 535–563 (1984).
524. Weis L., "The range of an operator in $C(X)$ and its representing stochastic kernel," *Arch. Math.*, **46**, 171–178 (1986).
525. Weispfenning V., "Model-completeness and elimination of quantifiers for subdirect products of structures," *J. Algebra*, **36**, No. 2, 252–277 (1975).
526. Westfall R., *Never at Rest. A Bibliography of Isaak Newton*, Cambridge University Press, Cambridge (1982).
527. Wolff M. P. H., "An introduction to nonstandard functional analysis," in: L. O. Arkeryd, N. J. Cutland, and C. Ward Henson (eds): *Nonstandard Analysis, Theory and Applications*, Kluwer Academic Publishers, Dordrecht, 1997, pp. 121–151.
528. Wolff M. P. H., "On the approximation of operators and the convergence of the spectra of approximants," in: *Operator Theory: Advances and Applications*, 1998, **103**, pp. 279–283.
529. Wright J. D. M., "Vector lattice measures on locally compact spaces," *Math. Z.*, **120**, No. 3, 193–203 (1971).
530. Yamaguchi J., "Boolean $[0, 1]$ -valued continuous operators," *Internat. J. Comput. Math.*, **68**, No. 1–2, 71–79 (1998).

-
531. Yessenin-Volpin A. S., "The ultra intuitionistic criticism and the traditional program for foundations of mathematics," in: *Intuitionism and Proof Theory*, North-Holland, Amsterdam and London, 1970, pp. 3–45.
 532. Yood B., *Banach Algebras. An Introduction*, Carleton University Press, Ottawa (1988).
 533. Young L., *Lectures on the Calculus of Variations and Optimal Control Theory*, Chelsea Publishing Co., New York (1980).
 534. Yushkevich A. P., "Leibniz and the foundation of infinitesimal calculus," *Uspekhi Mat. Nauk*, **3**, No. 1, 150–164 (1948).
 535. Zaanen A. C., *Riesz Spaces. Vol. 2*, North-Holland, Amsterdam etc. (1983).
 536. Zaanen A. C., *Introduction to Operator Theory in Riesz Spaces*, Springer-Verlag, Berlin etc. (1997).
 537. Zadeh L., *The Concept of a Linguistic Variable and Its Application to Approximate Reasoning*, Elsevier, New York etc. (1973).
 538. Zel'dovich Ya. B. and Sokolov D. D., "Fractals, similarity, and intermediate asymptotics," *Uspekhi Fiz. Nauk*, **146**, No. 3, 493–506 (1985).
 539. Zhang Jin-wen, "A unified treatment of fuzzy set theory and Boolean valued set theory fuzzy set structures and normal fuzzy set structures," *J. Math. Anal. Appl.*, **76**, No. 1, 297–301 (1980).
 540. Zivaljevic R., "Infinitesimals, microsimplexes and elementary homology theory," *Amer. Math. Monthly*, **93**, No. 7, 540–544 (1986).
 541. Zorich V. A., *Mathematical Analysis. Part 1* [in Russian], Nauka, Moscow (1981).
 542. Zvonkin A. K. and Shubin M. A., "Nonstandard analysis and singular perturbations of ordinary differential equations," *Uspekhi Mat. Nauk*, **39**, No. 2, 77–127 (1984).

Notation Index

- $\text{ltd}(\cdot)$, 18
 $\approx \mathbb{R}$, 18
 \ll , 18
 $\text{Fnc}(\cdot)$, 42
 $\text{Func}(\cdot)$, 42
 IST , 64
 St , 64
 $\varphi \in (\text{IST})$, 64
 $\varphi \in (\text{ZFC})$, 64
 \mathbf{V}^{st} , 65
 $(\forall^{\text{st}}x)\varphi$, 65
 $(\exists^{\text{st}}x)\varphi$, 65
 $(\forall^{\text{st fin}}x)\varphi$, 65
 $(\exists^{\text{st fin}}x)\varphi$, 65
 $^\circ x$, 65
 EXT , 73
 NST , 73
 $\text{Int}(\cdot)$, 74
 \mathbf{V}^{Ext} , 74
 \mathbf{V}^{Int} , 74
 \mathbf{V}^{St} , 74
 $\varphi \in (\text{EXT})$, 74
 V^{size} , 74
 φ^{St} , 74
 φ^{Int} , 74
 NCT , 94
 $\text{Psls}(D, p)$, 99
 $\lceil \varphi \rceil$, 101
 $\text{Su}(\cdot)$, 107
 $x \text{ st } y$, 108
 RIST , 112
 $\mu(\cdot)$, 117, 123, 124
 ${}^a\mathcal{B}$, 117
 (X, τ) , 123
 $h(G)$, 123
 $\text{nst}(G)$, 123
 $\text{cl}_{\approx}(U)$, 123
 $x \approx y$, 130
 $\approx A$, 130
 $\text{pst}(X)$, 134
 $x \overset{\tau}{\approx} a$, 140
 \mathcal{P}_{fin} , 143
 \mathcal{G}^\uparrow , 154
 \mathcal{F}^\downarrow , 154
 (x) , 157
 $e(X)$, 157
 A^c , 158
 $p(X)$, 158
 μ_d , 160
 $(\mathcal{F})^\uparrow\downarrow$, 162
 \mathcal{E}^\downarrow , 162
 $\text{cnt}(X)$, 167
 $\text{ltd}(X)$, 169
 B_X , 170, 228
 $\text{Ha}(F, x')$, 171
 $\text{Bo}(f, x')$, 171
 $\text{Cl}(F, x')$, 171
 $\text{H}(F, x')$, 171

- $\text{Fd}(F, x')$, 171
 $\text{K}(F, x')$, 171
 $\exists\exists\exists(F, x')$, 172
 $\forall\forall\forall(F, x')$, 173
 $\forall\exists\exists(F, x')$, 176
 $\forall\alpha\forall h\exists x(F, x')$, 177
 $\exists h\exists x\forall\alpha(F, x')$, 177
 $\exists h\forall x\forall\alpha(F, x')$, 177
 $\text{Ha}^+(F, x')$, 177
 $\text{R}^1(F, a')$, 178
 $\text{Q}^1(F, a')$, 178
 $\text{QR}^2(F, a')$, 178
 $0 = \{0\}$, 178
 $\text{R}^2(F, a')$, 178
 $\text{Q}^2(F, a')$, 178
 $\text{QR}^1(F, a')$, 178
 $\forall\forall(F)$, 184
 $\exists\forall(F)$, 184
 $\exists\exists(F)$, 184
 Li , 185
 Ls , 185
 $f(\text{Ha}_\alpha)(x')(h')$, 193
 $f(\text{In}_\alpha)(x')(h')$, 193
 $f(\text{Cl}_\alpha)(x')(h')$, 193
 Df , 206
 $\varepsilon_{\mathcal{A}}$, 209
 $\langle \mathcal{A} \rangle^e$, 209
 $\text{ltd}(E)$, 224
 $\mu(E)$, 224
 $E^\#$, 224
 $(\cdot)^\#$, 224
 $T^\#$, 225
 $B_L^0(a, \varepsilon)$, 225
 $B_L(a, \varepsilon)$, 225
 $\text{Seq}(f)$, 225
 $\dim(E)$, 226
 $l_p(n)$, 227
 $\sigma(T)$, 230
 $\sigma_p(T)$, 230
 $\text{DAp}(A)$, 233
 $\mathbf{t}(f) = T_N(f)^\#$, 234
 $A_N^\#$, 234
 $B^{(\lambda)}$, 235
 $R_\lambda(A)$, 238
 $S(\mathcal{A})$, 245
 ν_L , 245
 $(X, S_\Delta, \nu_\Delta)$, 248
 (X, S^X, ν_L^X) , 248
 S_Δ^M , 248
 ν_Δ^M , 248
 $(M, S_\Delta^M, \nu_\Delta^M)$, 248
 \mathcal{L} , 249
 \mathcal{S} , 249
 ν_Δ , 249
 \mathcal{L}_p , 251
 $\mathcal{L}_{p,\Delta}^X$, 251, 262
 $\|\cdot\|_{p,\Delta}$, 251
 $\|F\|_{p,A}$, 251
 $\mathcal{S}_p(M)$, 251
 $\|F\|_{p,\Delta}$, 262
 \mathcal{A}_k , 263
 A_K , 264
 $\varphi \otimes \psi$, 265
 $\mathcal{L}_{2\Delta}$, 282
 Φ_Δ , 282
 $\hat{\Delta}$, 282
 $X_\Delta(f)$, 282
 $X_\Delta(f)_k$, 282
 \mathcal{F} , 282, 295
 \mathcal{D}_d , 286
 \dagger , 287
 \div , 287
 $\mathcal{L}^{(n)}$, 288
 $\mathcal{L}^{(\sigma)}$, 288
 ψ_F^Δ , 288
 $C^{(\sigma)}$, 289
 $C^{(n)}$, 289
 \mathcal{M}_d , 290
 G_0 , 294, 316
 G_f , 294, 316

-
- | | |
|-------------------------------------------------------|----------------------------|
| \widehat{G} , 294 | \mathcal{F}_n , 341 |
| $\overset{\circ}{F}$, 296 | \mathcal{F}_n^{-1} , 341 |
| $\text{In}(G_f)$, 296 | $\mathcal{S}_p(G)$, 341 |
| $\text{In}_0(G_f)$, 296 | T_n , 341 |
| $G^\#$, 296, 316 | \widehat{T}_n , 341 |
| $S_\Delta^{G_f}$, 299 | $L_p(G_n)$, 341 |
| $G^\wedge := \widehat{G}$, 302 | $L_p(\widehat{G}_n)$, 341 |
| S^1 , 302, 328 | ι , 343 |
| $G^{\wedge\#}$, 302 | $\widehat{\iota}$, 343 |
| G'_0 , 303 | D_{test} , 344 |
| G'_f , 303 | \widehat{D}_{test} , 344 |
| $G^{\#\prime} := G'_f/G'_0$, 303 | S_n , 349 |
| $\mathcal{L}_{2,\widehat{\Delta}}(\widehat{G})$, 306 | \mathcal{X} , 352 |
| Φ_Δ^G , 306 | \mathbf{t} , 352 |
| $\mathcal{F}_\Delta^{G^\#}$, 306 | X_N , 352 |
| $\widetilde{\varphi}$, 307 | proxy($X_N^{(b)}$), 352 |
| $CS(G)$, 308 | A_f , 355 |
| \widetilde{j} , 316 | $A_f^{(n)}$, 355 |
| Ext, 318 | $S_n^{(2)}$, 356 |
| Δ_τ , 330 | $B_f^{(n)}$, 356 |
| $a \mid b$, 330 | $\widehat{\iota}$, 357 |
| $\text{rem}(a, b)$, 330 | W_f , 364 |
| $\mathbb{Q}^{(\tau)}$, 331 | $[\cdot]$, 380 |
| $\mathbb{Z}^{(\tau)}$, 331 | $\mathbf{1}$, 381 |
| Σ_τ , 332 | \models , 381 |
| \mathbb{Q}_p^+ , 335 | $\mathbb{V}^{(B)}$, 382 |
| \mathbb{Q}_α , 338 | $X\downarrow$, 382 |
| | $X\uparrow$, 383 |

Subject Index

- *-image, 79
- \mathcal{A} -measurable internal function, 245
- B -structure, 383
- n -ary symbol, 39
- n -place symbol, 39
- p -monad, 98
- p -saturated class, 99
- p -section, 95
- p -standard class, 95
- Q_1Q_2 -limit, 184
- S -continuous function, 307
- S -continuous representation, 313
- \mathcal{S} -integrable internal function, 246
- \mathcal{S}_M -integrable function, 248
- \mathcal{S}_p -integrable lifting, 301
- $\mathcal{S}_{p,\Delta}$ -lifting, 301
- U -net, 136
- κ -concurrency, 82
- κ -saturation, 83
- σ -finite subspace, 248
- σ -permutation, 89
- Σ_0 -formula, 45
- Σ_1 -formula, 45
- τ -adic solenoid, 332
- τ -infinitesimal, 112, 143
- τ -monad, 140
- τ -random element, 253
- τ -strongly-standard element, 114
- ε -subdifferential, 205
- ε - δ -technique, 7
- abstract Nemytskiĭ operator, 373
- abstract relation, 42
- abstract Urysohn operator, 373
- adjoint, 230
- admissible triple, 306
- alephic scale, 52
- Alexandrov, 13
- alphabet, 39
- antisymmetric relation, 45
- appropriate size, 78
- approximable group, 340
- approximating sequence, 325
- approximation domain, 233
- Archimedes, 5
- ascending and descending, vii
- ascent, 383
- ascent of a filter, 154
- assignment operator, 40
- astray element, 118
- at most countable set, 53
- atomic formula, 39
- Attouch, 222
- Auerbach basis, 229
- autohalo, 123
- auxiliary symbols, 39
- axiom of acceptance, 78
- axiom of choice, 49, 89
- axiom of complement, 88
- axiom of comprehension, 48, 88
- axiom of domain, 89
- axiom of embedding, 75
- axiom of extensionality, 47, 87
- axiom of foundation, 49
- axiom of infinity, 49, 87
- axiom of intersection, 88
- axiom of membership, 88
- axiom of pairing, 48, 87
- axiom of powers, 48

- axiom of powerset, 48, 87
 axiom of product, 89
 axiom of regularity, 49, 89
 axiom of replacement, 48, 88
 axiom of separation, 48
 axiom of the empty set, 88
 axiom of transitivity, 75
 axiom of union, 47, 87
 axioms of interplay, 74
 axioms of permutation, 89
- basic field, 224
 Berkeley, 5, 6
 Bernays–Morse theory, 93
 Bernoulli, 2
 Bernoulli automorphism, 254
 binary relation, 43
 Bolzano, 7
 boolean, 43
 Boolean algebra of countable type, 256
 Boolean valued analysis, vii
 Boolean valued model, 382
 Boolean valued universe, 382
 Bouligand, 170, 171, 176
 Bouligand cone, 170
 bound occurrence, 40
 bound variable, 39, 86
 bounded formula, 45
 bounded formulas are absolute, 76
 Bounded Idealization Principle, 72
 bounded point, 169
 bounded quantifier, 45
 bounded set, 72
 Boundedness Principle, 72
- canonical embedding, 382
 canonical enlargement, 57
 canonical Loeb measure space, 248
 Cantor, 10, 12, 36, 37
 cantorion set, 12
 cardinal, 52
 cardinal number of a set, 53
 cardinality, 46, 52, 53
 Carnot, 7
 cartesian product, 42
 Cauchy, 7, 180, 183
 Cauchy principle, 84, 144
 centered family, 82
 chain, 45
 characteristic, 256
 Church schema, 47
 Clarke, 170, 176, 222
 Clarke cone, 170
 Clarke derivative, 193
- class, 47, 64, 86
 class of all sets, 47
 class-function, 42
 classical universe, 79
 classification, 47
 cofinite filter, 159
 compact correspondence, 150
 compact filter, 148
 compact net, 148
 compact sequence of operators, 241
 compactum, 129
 compatible sequential approximant, 345
 complete filter, 149
 complexity of a formula, 91
 composite, 44
 composition, 44
 conatus of a space, 167
 conatus of a vector, 167
 concurrent correspondence, 82
 condition (*), 297
 cone of feasible directions, 171
 conservative theory, 93
 containment, 40
 contingency, 171
 continuum, 50
 continuum hypothesis, 53
 continuum problem, 53
 Cornet Theorem, 196
 correspondence, 43
 countable chain condition, 256
 countable set, 53
 counting measure with weight, 248
 Courant, 24
 cumulative hierarchy, 51
 cyclic complement, 158
 cyclic filterbase, 154
 cyclic grill, 162
 cyclic monad, 156
 cyclic monadic hull, 160
 cyclic set, 382
 cyclically compact space, 159
- D’Alembert, 6
 Dacunha-Castelle, 232
 de l’Hôpital, 2
 definor, 40
 derivatives along effective domain, 193
 descent, 382
 descent of a family, 162
 descent of a filter, 154
 deviation, 131
 differential, 2
 direction, 167, 182
 directionally Lipschitz mapping, 195

- discrete approximation, 233
 discrete compactness, 241
 discrete convergence, 233
 discrete filter, 184
 discrete Fourier transform, 282, 306
 discrete spectrum, 235
 disjointly additive operator, 373
 distant element, 117, 118
 Dolecki, 222
 domain, 42, 43
 domain of arrival, 43
 domain of definition, 43
 domain of departure, 43
 dual hyperapproximant, 321
 dual sequential approximant, 325

 eigenfunction, 310
 eigenvalue, 310
 element, 40
 element of quantity, 8
 empty class, 88
 empty set, 41
 epi-Lipschitz mapping, 171
 epilimits, 188
 equality sign, 39
 equipollent sets, 46, 52
 equipotent sets, 46, 52
 equivalent nets, 182
 ergodic operator, 254
 essential point, 157
 Euler, 5, 17, 19, 20, 27, 32, 34, 35, 376
 Euler formula, 5
 expression, 38
 extended reals, 186
 extension of a group, 318
 extension principle, 83
 extensional correspondence, 383
 extensional filterbase, 154
 external extension, 60
 external formula, 64
 external quantifier, 69
 external set, 64
 external set formation, 74
 external universe, 74

 falsity, 40
 Fenchel, 207, 211
 Feuerbach, 73
 filtered formula, 56
 filtered power, 56
 finite Loeb measure space, 246
 first-order language, 38
 first-order theory, 40
 formal system, 38

 formula of signature σ , 39
 free occurrence, 39
 free variable, 39
 function, 43
 function symbol, 39
 function-class, 86, 89
 function-set, 42

 generalization, 86
 generalized continuum hypothesis, 53
 generalized point, 210
 generalized problem of the continuum, 53
 Gödel, ix
 Gordon, 377
 granted horizon principle, 122
 graph, 43

 Hadamard, 170, 171, 176
 Hadamard cone, 170
 halo, 123
 Hartogs number, 52
 hermitian conjugate, 230
 Hilbert–Schmidt operator, 264
 Hiriart-Urruty, 222
 Hrbáček, viii, 73
 Hrbáček Theorem, 76
 hyperapproximant to a group, 316
 hyperapproximant to a measure space, 259
 hyperapproximant to an operator, 262
 hyperapproximation of an operator, 262
 hyperconvexity, 168
 hyperfinite realization of a measure space, 259
 hyperfinite-dimensional vector space, 225, 226
 hyperrepresentation, 313
 hypertangent cone, 171

 idea of variability, 9
 idealization principle, 65, 76
 identity mapping, 44
 identity relation, 44
 Π' in, 11
 image, 42
 image of a set, 43
 independent sequence, 254
 indivisible element, 3
 inductive limit, 60
 infinite cardinal, 52
 infinite closeness, 205
 infinite proximity, 129, 205
 infinitely close points, 130
 infinitely large element, 118, 182
 infinitely small entity, 4
 infinitely small neighborhood, 123

- infinitely small quantity, 9
 infinitely small set, 133
 infinitesimal element, 117
 infinitesimal gradient at a generalized point, 210
 infinitesimal in a normed space, 224
 infinitesimal limit test, 23
 infinitesimal neighborhood, 123
 infinitesimal Pareto solution, 221
 infinitesimal set, 133
 infinitesimal solution, 219
 infinitesimal subdifferential, 206, 211
 infinitesimal subgradient, 206
 instantiation, 86
 internal class, 64, 95
 internal dimension, 226
 integral operator, 263
 internal quantifier, 69
 internal saturation principle, 72
 internal set, 64
 internal set technique, 83
 internal universe, 74
 internalization, 74
 internally linearly independent set, 226
 inverse of a correspondence, 44
 inverse relation, 44

 Kachurovskii, 378
 Kawai, viii, 73
 Kawai Theorem, 79
 Keldysh Theorem, 371
 killing of quantifiers, 166
 Kolmogorov, 170
 Krivine, 232
 Kuratowski, 37, 180, 185, 186, 196, 199
 Kuratowski limit, 185
 Kuratowski lower limit, 184
 Kuratowski upper limit, 184
 Kuratowski–Zorn lemma, 49

 Landau, 20
 language of set theory, 40
 Leibniz, 2, 4, 5, 20, 31, 36
 Leibniz principle, 80, 83
 length of a formula, 91
 lifting, 245, 249
 limit inferior, 184, 188
 limit ordinal, 50
 limit superior, 184, 188
 limited element in a normed space, 224
 limited part of a space, 169
 limited point, 169
 limited set, 72
 limitedness principle, 44, 184

 linear order, 45
 Lipschitz, 170, 171, 195
 Loeb measure, 245
 Loeb measure space, 245
 Loeb-measurability, 245
 Łoś Theorem, 56
 lower limit, 188
 Luxemburg, 159, 232, 367
 Luzin, x, 8, 9, 24, 36

 Maharam, 209
 Mal'tsev, 9
 mapping, 42
 mathematical analysis, 4
 maximality principle, 49
 measure, 256
 member, 40
 membership, 40
 metalanguage, 38
 metasymbol, 41
 method of indivisibles, 4
 method of prime and ultimate ratios, 3
 Michael Theorem, 375
 microclosure, 123
 microcontinuous function, 130
 microcontinuous function at a point, 30
 microhalo, 130
 microlimit point, 123
 Micromegas, 6
 microstep mapping, 138
 mixing, 382
 modeling principle, 75
 monad, 3, 18, 117, 124
 monad of a set, 98
 monad of a topological vector space, 167
 monadic hull, 160
 Montague–Levy Theorem, 53
 Moore subnet, 182
 Mostowski, 37
 multimetric, 131
 multimetric space, 131

 naive set, 12
 naive standard set, 12
 nearstandard part, 123
 nearstandard point, 127
 nearstep function, 137
 neartopological vector space, 168
 nearvector topology, 168
 negligible set, 252
 Nelson, viii, 180, 181
 Nelson algorithm, 169
 neoclassical stance, 84
 nest, 98

- Newton, 3, 4, 6
 nice hyperapproximant, 317
 nonstandard analysis, vii
 nonstandard element, 14
 nonstandard hull, 224, 279
 nonstandard hull of an operator, 225
 normalizing factor, 299
 normed Boolean algebra, 256
- operation symbol, 39
 operator norm, 224
 order, 45
 ordered n -tuple, 42
 ordered pair, 42
 ordinal, 50
 ordinal rank, 51
 orthogonality in a normed space, 229
 overflow principle, 83
- Pareto, 221
 partition, 243
 permanence principle, 84, 144
 Planck constant, 292
 Pliev, 371
 point spectrum, 230
 polar, 44
 polyadic integers, 339
 Pontryagin, 1, 37
 Powell Theorem, 66
 powerset, 43
 predicate calculus, 40
 predicate symbol, 39
 predicative formula, 91
 predicative formula of NCT, 95
 prenearstandard point, 134
 preorder, 45
 pretopological space, 123
 principle of induction on rank, 51
 principles of infinitesimal analysis, 65
 principles of nonstandard set theory, 65
 pro-total-boundedness, 159
 procompact space, 159
 procomplement, 158
 progrill, 162
 proideal point, 158
 projective limit, 329
 proper class, 86
 property, 47
 proposition, 40
 propositional connectives, 39
 protoextension, 58
 poultrafilter, 154
 proxy-standard element, 352
 pseudodifferential operator, 355
- quantifiers, 39
 quasicompact sequence of operators, 236
- radical stance, 85
 random element, 252
 random finitely additive measure, 277
 random measure, 277, 369
 range, 42
 rank, 51
 rapid decay, 325
 rapidly decreasing function, 325
 record, 38
 reflection principle, 53
 reflexive relation, 45
 regularizing cones, 177
 relation, 43
 relatively standard set, 376
 relativization, 55, 74
 remote element, 117, 118, 182
 representative set, 192
 Reshetnyak, 11
 resolution of the identity, 256
 restriction, 43
 ring of τ -adic integers, 330
 Robinson, 4, 6, 9, 70, 159
 Robinson Lemma, 70
 Robinson principle, 84, 144
 robinsonian standardization, 79
 Rockafellar, 180, 188, 190, 193, 195, 222
 Rockafellar derivative, 193
 Rockafellar formula, 190
 Rockafellar limit, 188
 Rokhlin–Kuratowski–Ryll–Nardzewski
 Theorem, 375
- Sadovnichii, 11
 saturated set, 123
 saturation principle, 83
 Schrödinger-type operator, 360
 Schwartz, 10
 Scott, 384
 semimetric, 131
 semimetric space, 131
 semiset, 99
 Sendov, 11
 sentence, 40
 sequential approximant, 324, 340
 sequential normalizing factor, 341
 set, 51, 86, 94
 set of standard size, 74
 Severi, 7
 shadow, 378
 signature, 39
 signature of a language, 39

- simple internal function, 246
- single-valued correspondence, 42
- singleton, 49
- size, 16
- Solovay, 384
- son, 50
- spectrum, 230
- standard core, 65
- standard environment, 14, 71
- standard part, 65, 77
- standard part operation, 23
- standard size, 106
- standard universe, 74
- standardization of a formula, 74
- standardization principle, 65, 76
- standardly finite set, 97, 297
- standardness, 64
- statement, 40
- strict subnet, 182
- strong concurrence principle, 82
- strong discrete approximation, 233
- strong uniformity, 132
- subcontinuous correspondence, 150
- subdifferential calculus, 166
- subdirection, 182
- subnet, 182
- subnet in a broader sense, 182
- subordinate net, 183
- subuniverse, 57
- successor, 50
- superposition, 44
- Swift, 6
- symbol of an operator, 355
- symmetric relation, 45

- table, 263
- Takeuti, 384
- target, 43
- term, 39
- text, 38
- theorem, 38
- Thibault, 222
- tiny finite cover, 138
- topological space, 123
- total order, 45
- totally bounded filter, 149
- transfer principle, 65, 76
- transitive relation, 45, 50
- triple, 90

- ultralimit construction, 58
- ultranet, 183
- ultrapower, 56
- ultrapower construction, 55

- unbound variable, 39
- uncertainty principle, 292
- underflow principle, 83
- uniform compactness, 241
- uniform discrete convergence, 233
- uniform Loeb measure, 248
- uniformity, 129
- uniformity of pointwise convergence, 132
- unique determination principle, 71
- unit ball, 228
- universal class, 88
- universal closure, 40
- universal mathematics, 3
- universe, 47
- universe of sets, 12
- universet, 55
- universoid, 55
- unlimited real, 17
- unordered pair, 41
- upper limit, 188

- vanishing quantity, 3
- variables, 39
- Varignon, 4
- verity, 40
- Vietoris uniformity, 132
- Voltaire, 6
- von Neumann universe, 49
- von Neumann–Gödel–Bernays theory, 86
- Vopěnka, 37, 384
- Vygodskii, 8

- Ward, 222
- weak convergence, 132
- weak concurrence principle, 82
- weak idealization principle, 57
- Weierstrass, 7
- weight of a group topology, 327
- well- τ -saturated set, 123
- well-formed formula, 38
- well-orderable set, 45
- well-ordered set, 45, 50
- well-ordering principle, 49
- well-saturated set, 123
- Weyl symbol, 364

- Young, 37, 207, 211

- Zermelo set theory, 75
- Zermelo subset, 57
- Zermelo Theorem, 49, 75
- Zermelo universet, 57
- Zermelo–Fraenkel set theory, 40, 47
- Zorich, 11, 16