Universität Bielefeld

Ultraproduct Construction of Representative Utility Functions with Infinite-Dimensional Domain

Dissertation zur Erlangung des akademischen Grades Dr. math.
der Fakultät für Mathematik
der Universität Bielefeld
vorgelegt von

Geghard Bedrosian

November 2015

Erstbetreuer: Prof. Dr. Frank Riedel
Zweitbetreuer: Prof. Dr. Dr. Frederik Herzberg

Bielefeld University

Ultraproduct Construction of Representative Utility Functions with Infinite-Dimensional Domain

A thesis submitted in fulfilment of the requirements for the degree of Dr. math.
at the Department of Mathematics
Bielefeld University
Submitted by

Geghard Bedrosian

November 2015

Supervisors:
Prof. Dr. Frank Riedel
Prof. Dr. Dr. Frederik Herzberg

Gedruckt auf alterungsbeständigem Papier nach DIN-ISO 9706

Acknowledgements

Foremost, I would like to thank my supervisors Prof. Frank Riedel and Prof. Frederik Herzberg for providing me with the opportunity to complete my thesis. I would like to appreciate the support of Prof. Frank Riedel who accepted me as a Doctoral student in his group. I especially want to thank Prof. Frederik Herzberg for his valuable guidance, scholarly inputs and consistent encouragement I received throughout this research work. This feat was possible only because of the unconditional support provided by him. A person with an amicable and positive disposition who has always made himself available to clarify my doubts despite his busy schedules. I consider it as a great opportunity to do my Doctoral program under his guidance and to learn from his research expertise.

Special thanks are also given to Prof. Alessandra Palmigiano. She is the one who invited me as a visiting scholar to TU Delft University and for the last stage of my dissertation, she suggested a very interesting project and helped and encouraged me to overcome any difficulty I encountered in this way. During my stay in Delft, I met Zhiguang Zhao who I think is one of the best mathematical logicians and got this possibility to exchange a lot of research ideas with him. Thank you Zhiguang. I am very grateful to Prof. James Hartley, Prof. Carlos Hervés Beloso, Prof. Daniel Eckert, Dr. Patrick Beißner, Dr. Oktay Sürücü and Dr. Giorgio Ferrari for helpful comments and suggestions. My appreciation also goes to Prof. Thomas Zink.

I gratefully acknowledge financial support by the German Research Foundation (DFG) through the International Graduate College (IGK) *Stochastics and Real World Models* (Bielefeld–Beijing), as well as the support of the Chinese part of the IGK during our stays in China. My gratitude goes to Prof. Michael Röckner (Spokesperson of the graduate college), all the IGK and IMW (Center for Mathematical Economics) professors for the insightful lectures. I would like to thank Rebecca Reischuk, Karin Zelmer and Hanne Litschewsky for their administrative supports. I am grateful to my colleagues from IGK and IMW for their supports, especially to Diana (for reading and correction of my thesis), Jesper, Andrea, Tolu, Viktor, Mykola, Milan, Yuhua, Marina and Tobias.

I am particularly indebted to my family and can not find enough words to express how grateful I am to my parents. They have encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true. I miss them deeply.

Last but not least, my hearty thanks go to my devoted and beloved fiancée, Jizet. She forms the backbone and origin of my happiness. Her love and support without any complaint or regret has enabled me to complete this Doctoral project. I am also indebted to Jizet's family, especially to her father, mother and brother for their heartfelt supports. Thank you for everything.

To Jizet

Contents

Chapter 1

Introduction

Modern macroeconomic theory looks for microeconomic foundations, namely consumers and firms, and it considers how these microeconomic entities in an economy make their decisions and then how these many individuals' choices give rise to economy-wide macroeconomic outcomes (see Gillman [23] and Mankiw [43], [44]). The earliest macroeconomic models were based on hypotheses about relationships between aggregate quantities, such as aggregate output, employment, consumption, and investment. Weintraub [63] mentioned that critics and proponents of these models disagreed as to whether these aggregate relationships were consistent with the principles of microeconomics. Böhm-Bawerk in 1895 was possibly the first economist to insist that propositions about macroeconomics have firm microeconomic foundations. In [13] (see also page 74 in Hennings [28]), he wrote:

> *"One cannot eschew studying the microcosm if one wants to understand properly the macrocosm of a developed economy."*

Therefore, in recent decades macroeconomists have attempted to combine microeconomic models of consumer and firm behaviour to derive the relationships between macroeconomic variables, and a revitalised group of new macroeconomic theorists prefer to adapt micro to macro theory for explaining economic fluctuations. Important figures here include Akerlof [1], Blanchard [11], Gordon [25], Mankiw and Romer [45] and Stiglitz [59].

> *"All macroeconomic events are the accumulation of millions of decisions made by individual people. From this observation, it seems to follow immediately that if we are ever going to understand macroeconomy, we need to understand the microeconomic behaviour which forms its basis."*
> -James Hartley

Since in realistic models consumers and firms are heterogeneous, an accurate and comprehensive mathematical description of the aggregate behaviour is typically an intractable problem. One solution is to assume the existence of a *representative agent* in a loosely defined sense. Indeed, in mathematical models of macroeconomic theory, the assumption of a representative agent is ubiquitous (see Snowdon and Vane [58], Mankiw [42] and Sargent [56]), although it is controversial. The notion of the representative agent can be traced back to the 20th century. Marshall [47] in 1920 introduced a representative firm, however, after Lucas' critique in 1976 which motivated development of microfoundations for macroeconomics, the notion of the representative agent became more prominent and more controversial. Lucas [39] pointed out that policy recommendations based on observed past macroeconomic relationships may neglect subsequent behavioural changes by economic agents, which, when added up, would change the macroeconomic relationships themselves. The representative agent models are thus an attempt to model rigorously the structural relationships in an economy and for this purpose, one needs to model the behaviour of individuals, which is exactly the same as saying one needs to provide microfoundations for the macroeconomic models.

Hartley [26] (see also Hartley [27]) found this reason for the representative agent model unconvincing and he punctuated that the representative agent models do not meet the goals for which they are constructed. Later on Kirman [34] provided an example in which the representative agent disagrees with all individuals in the economy. He concluded that the reduction of a group of heterogeneous agents to a representative agent is not just an analytical convenience, but is both unjustified and leads to conclusions which are usually misleading and often wrong.

The search for a rigorous justification for the representative agent models has so far been unsuccessful and was ultimately abandoned until very recently. Clark [16] developed a theory of the representative agent, whose optimal behaviour is equivalent to the collective behaviour of a society of rational individuals. He proposed a method for aggregating the random preference relation into a determinate preference relation on random actions[1]. Later on, Herzberg [31] constructed a representative utility function for finite-dimensional social decision problems, based on a bounded ultrapower construction over the real numbers, with respect to the ultrafilter induced by the underlying social choice function (via the Kirman-Sondermann [35] correspondence).

However, since the decision problems of macroeconomic theory are typically infinite-dimensional[2], Herzberg's original result is insufficient for many applications. Therefore, *one of the main contributions of our work is to generalise his result to the case of*

[1]There is some literature about the construction of a representative agent in general equilibrium theory (for example see Negishi [48]) which does not provide a social choice theoretic foundation and for our work, we are not interested in this kind of representative agents.

[2]Macroeconomic studies emphasise decisions with a time dimension and in practice, the macroeconomic framework used in research almost always involves an infinite number of time periods (see Aliprantis et al. [3] and Ljungqvist and Sargent [37]). The reason for this is mainly analytical convenience: it allows stationary in the sense that the remaining time horizon is the same – it is infinite – as time passes. It is essential that the time horizon be long enough that the idea that "the world ends soon" not influence the results too much.

infinite-dimensional social decision problems and to pave the way for applications in actual macroeconomic models.

Mathematical framework

For achieving our goal in this work, we combine arguments from mathematical logic and analysis with arguments from social choice theory. The theory of social choice is the formal study of mechanisms for collective decision making, and investigates issues of philosophical, economic, and political significance, stemming from the classical Arrovian problem of how the preferences of the members of a group can be democratically aggregated into one outcome.

In the last decades, *ultrafilters* and *ultraproducts* from mathematical logic have facilitated the build-up of a critical mass of results and insights generalising the original Arrovian problem. Arrow [7] showed that if the set of individuals is finite and the set of alternatives is at least three, then Arrow's axioms[3] which are intuitive and reasonable, are inconsistent. This theorem is called *Arrow's Impossibility Theorem.* In the same logical vein, Fishburn [21] showed that if the set of individuals is infinite and the set of alternatives is at least three, then Arrow's axioms are consistent, which is called *Fishburn's Possibility Theorem.* Kirman and Sondermann [35] proved that there is a one to one correspondence between the social welfare function satisfying all Arrovian rationality axioms and the set of decisive coalitions generated by this social welfare function, which is an ultrafilter on individuals. Armstrong [5], [6] generalised Kirman and Sondermann's results to measure spaces of individuals. For the first time, Lauwers and Van Liedekerke [36] used ultraproducts to show that there is a one to one correspondence between the aggregation function, which satisfies all Arrovian rationality axioms, and non-principal ultrafilters. Herzberg and Eckert [32] generalised Kirman and Sondermann's correspondence by using model-theoretic method. They characterised Arrovian rationality social welfare functions as being exactly those defined with respect to an ultraproduct construction parametrised by the ultrafilter induced by the associated decisive coalitions. Ulam [61] and Tarski [60] proved the existence of non-principal ultrafilters under the assumption of the Axiom of Choice and Łoś [38] proved Łoś's Theorem which is one of

[3]Arrow's axioms are:

1. Unrestricted domain (universality): All preferences of all individuals (votes) are allowed.
2. Unanimity preservation: If every individual prefers alternative x over alternative y, then the society prefers x over y.
3. Independence of irrelevant alternatives: If every individual's preference between x and y remains unchanged, then the society's preference between x and y will also remain unchanged (even if individuals' preferences between other pairs like x and z, y and z, or z and v change).
4. No dictatorship: No single individual possesses the power to always determine the society's preference.

the main results on ultraproducts. Makkai [41] replaced the standard model-theoretic ultraproduct construction with a generalised one in a category-theoretic setting.

We employ a very special ultraproduct construction, the so called *bounded ultrapowers*, which are used to build foundations for *nonstandard analysis* (see Robinson and Zakon [53]). In other words, we apply a bounded ultrapower with respect to a given ultrafilter as a framework for nonstandard analysis.

> *"There are good reasons to believe that nonstandard analysis, in some version or other, will be the analysis of the future."*
> -Kurt Gödel

Nonstandard analysis was established for the development of calculus and of other branches of analysis in terms of infinitely small and infinitely large quantities. The notion of an infinitesimal (infinitely small) number has been used in mathematical arguments since before the time of Archimedes (see Dijksterhuis [19]). In the late 1600's, Newton and Leibniz, in particular Leibniz used infinitesimal numbers to define the derivative and the integral (see Baron and Bos [8] and Child [15]). He created the notion dx for the difference in successive values of a variable x, thinking of this difference as infinitely small or less than any assignable quantity. Loeb in [64] said that it is helpful to think in terms of infinitesimals in branches of mathematics beyond calculus. He continued with this example: Brownian motion is the random motion of microscopic particles suspended in a liquid or gas. Under a microscope you can clearly see the zig-zag random motion caused by the collision of the particles with molecules in the suspending medium. An important part of mathematical probability is the construction of mathematical models for this motion. A very convenient model has each particle performing a random walk with steps of infinitesimal length. That is, one divides time into infinitesimal intervals, and in each interval, the particle moves in a straight line over an infinitesimal distance equal to the square root of the time change. At the end of each time interval, the particle chooses at random a new direction for its motion.

> *"In the fall of 1960 it occurred to me that the concepts and methods of contemporary mathematical logic are capable of providing a suitable framework for the development of the differential and integral calculus by means of infinitely small and infinitely large numbers."*
> -Abraham Robinson

Nonstandard analysis was originated in the early 1960's by the mathematical logician Abraham Robinson. He gave a clear, mathematically correct foundation for the use of infinitesimals in all branches of mathematics (see Robinson [51]). He set up a formal language to express facts about the mathematical objects with which he was working. He called this mathematical object a standard model for the theorems expressed in this formal language. He also showed that there is another mathematical object, called a nonstandard model, in which there exist positive infinitesimal numbers. The most important achievement of his work was that he showed any theorem which is a correct statement in the standard model, is also a correct statement in the nonstandard model (Transfer Principle). In 1955, Łoś [38] proved the Transfer Principle for any hyperreal number system. Its most common use is in Robinson's nonstandard analysis of the hyperreal numbers, where the Transfer Principle states that any sentence expressible in a certain formal language that is true of real numbers is also true of hyperreal numbers. Furthermore, Robinson turned the Transfer Principle into a working tool of mathematical reasoning. In the last few decades it has been applied to many areas, including analysis, topology, algebra, number theory, mathematical physics, probability and stochastic processes, and mathematical economics[4].

In addition to many mathematical techniques that we use in this work, like convex analysis, optimisation problems, functional analysis and etc., there are some reasons that we choose mathematical logic and nonstandard analysis as a main technique for our work. First of all, so far except Herzberg [31], neither bounded ultrapowers nor nonstandard analysis have ever been applied in the social choice theory literatures. Secondly, nonstandard analysis offers some features to our understanding of mathematics. For example, by learning nonstandard analysis, we learn new definitions of familiar concepts, often simpler and more intuitively, new and insightful constructions of familiar objects as well as new and insightful (often simpler) proofs of familiar theorems.

The use of mathematical logic in our work is in some ways similar to the use of geometric language in mathematics. The intuition enabling us to read a "meaning" into a "sentence" of a formal language can be compared to the intuition that enables us to understand and perceive geometric truths by checking a diagram.

Structure of the thesis

As we have mentioned above, this thesis tries to give an answer to the following question:

> *Is it possible to generalise Herzberg's result [31] to the case of infinite-dimensional social decision problems and to pave the way for applications in actual macroeconomic models?*

In order to do this, we suppose that individuals have cardinal utilities, i.e. for every individual, there is an utility function which induces his (or her) preference ordering.

[4]One can find some literature for applications of nonstandard analysis in diverse areas. For example, Anderson [4] provides an introduction to nonstandard analysis with applications to mathematical economics. We suggest also Albeverio et al. [2] for nonstandard methods in stochastic analysis and mathematical physics, Salbany and Todorov [55] for nonstandard analysis in topology, Jin [33] for applications of nonstandard analysis in additive number theory, Wolf and Loeb [64] for applications of nonstandard analysis in topology and measure theory.

We suppose that every individual's utility function belongs to some class of functions, where the elements of this class are called *admissible utility functions*. Furthermore, the aggregation of individual preferences is assumed to result from a social welfare function satisfying all Arrovian rationality axioms (universality, unanimity preservation, independence of irrelevant alternatives and no dictatorship). In Theorem 3.17 which has been shown for the first time by Kirman and Sondermann [35], we show that the collection of decisive coalitions generated by a social welfare function satisfying all Arrovian rationality axioms is always a non-principal ultrafilter. This is only possible if the set of individuals is infinite.

We generalise Herzberg's result by allowing the set of social alternatives to belong to a general Banach space W. We suppose that this set is a compact non-empty convex subset of the given Banach space W (Herzberg [31] supposed that the set of alternatives is a finite-dimensional vector space). Furthermore, we assume that the set of admissible utility functions are parametrised and the parameter set is a compact subset of a given Banach space X. For our results, the set of parametrised admissible utility functions contains only continuous and strictly concave functions.

The key for existence of a representative utility function is the proof of existence of a *socially acceptable utility function*. A socially acceptable utility function is an admissible utility function whose maximiser can never be quashed in the ultrafilter hierarchy (in other words, the maximiser can never be quashed by any other decisive coalition). We call an admissible utility function *representative* if and only if the maximiser of this representative utility function is the optimal alternative according to the social preference relation.

Using a nonstandard enlargement of the superstructure over $(X \oplus W) \cup \mathbb{R}$, obtained by a bounded ultrapower construction with respect to the non-principal ultrafilter, we prove that there exists for every utility profile, some $\mathcal{D}$-socially acceptable utility function (Theorem 3.23). For the proof of this theorem[5], we introduce a superstructure over $(X \oplus W) \cup \mathbb{R}$, i.e. $V((X \oplus W) \cup \mathbb{R})$, where $X \oplus W$ is a Banach space given as the algebraic direct sum of two Banach spaces X and W with norm $\| x \oplus w \|_\infty = \max\{\| x\|_X, \| w\|_W\}$.[6] We define the superstructure $V((X \oplus W) \cup \mathbb{R})$ by iterating the power-set operator countably many times. Then we construct a bounded ultrapower of $V((X \oplus W) \cup \mathbb{R})$ by collecting the equivalence classes of sequences in $V((X \oplus W) \cup \mathbb{R})$, that are bounded in the superstructure hierarchy, using the non-principal ultrafilter $\mathcal{D}$ on N. Afterwards we embed this bounded ultrapower into the superstructure $V({}^*(X \oplus W) \cup {}^*\mathbb{R})$ in such a way that this embedding satisfies the *Extension* and *Transfer Principles*. We work on

[5]This is an informal discussion of the proof methodology. The complete proof can be found in Section 3.4.

[6]One can show that the two norms $\| x \oplus w\|_p = (\| x\|_X^p + \| w\|_W^p)^{1/p}$ for $1 \le p < \infty$ and $\| x \oplus w\|_\infty = \max\{\| x\|_X, \| w\|_W\}$ for $p = \infty$, are equivalent on $\mathbb{R}$.

this nonstandard universe, which consists of equivalence classes of sequences of superstructure elements with respect to the equivalence relation of "almost sure agreement" according to $\mathcal{D}$. On the *image of $X \oplus W$, a standard operator with respect to the canonical topology on $X \oplus W$ is definable. We verify that the standard part of the $\mathcal{D}$-equivalence class of a utility profile is the parameter of a $\mathcal{D}$-socially acceptable utility function. The proof of this assures the continuity and S-continuity arguments with features of the bounded ultrapower construction.

We use Theorem 3.23 for proving the existence of a σ-representative utility function (Theorem 3.24). These results depend on certain regularity features of the admissible utility functions[7] (Assumption 3.19) and we show that the maximiser of this representative utility function maximises the $\mathcal{D}$-socially acceptable utility function, where $\mathcal{D} := F_\sigma$ is the ultrafilter of σ-decisive coalitions. In other words, the maximiser of this representative utility function is the optimal alternative according to the social preference relation. The complete proof can be found in Section 3.4. For paving the way for applications in actual macroeconomic models, we provide sufficient conditions for the preceding theorems (Theorems 3.23 and 3.24) to be satisfied in economic applications. The idea is that, since the agents maximise some function (we call it *happiness function*) subject to the constraints they face, we consider that the utility function of each agent is the maximum value of this happiness function subject to some budget constraint, where the social alternative is given. Then we show that this utility function has the same properties as the utility functions in Assumption 3.19 (see Lemma 4.1). We prove that for these new utility profiles, there exists some $\mathcal{D}$-socially acceptable and σ-representative utility functions. Afterwards, in Section 4.2, we give an example of possible macroeconomic applications.

After these arguments, the following question comes to our mind:

> *Are the previous conditions and assumptions weak enough? Or could we generalise our results to the case of weak topology?*

For this purpose, we generalise and extend our previous results to the case of weak compact social decision problems. At first, we introduce a new and simple nonstandard account of the weak topology and provide a mathematical foundation for the application of the powerful tools of nonstandard analysis to the weak topology. We give a nonstandard universe for pseudo-normed linear spaces and study the properties of the nonstandard hull with respect to this new nonstandard universe (see Section 5.1.1 and Proposition 5.9). The notion of nonstandard hull was introduced for the first time by

[7] According to this regularity on the admissible utility functions, the representative utility function is considered as the utility function of some individual. Kirman and sondermann [35] studied the case where the society's preferences will be controlled by a group of individuals. They created an agent whose preferences represent those of the group and called this agent an *invisible agent*. To see this, they considered the case where the set of social alternatives is finite, which is not the case in our work.

Luxemburg in [40]. He considered the case when X is a linear normed space. Later, Henson and Moore in [29] introduced this notion for Banach spaces. But, we study the nonstandard hull for the pseudo-norm linear spaces.

Then we define a weak topology on a given normed linear space by some pseudo-norms on this space and investigate the properties of the nonstandard hull with respect to this new weak topology (see Section 5.1.2 and Proposition 5.18). We assume that the set of social alternatives forms a weakly compact non-empty convex subset of a given reflexive separable Banach space W. We also assume that the admissible utility functions are parametrised and the parameter set is a weakly compact non-empty subset of a given separable Banach space X. For our results, the set of parametrised admissible utility functions contains w-uniformly continuous, Gâteaux-differentiable with continuous derivative, and strictly concave functions. Due to the nice results that we achieve about weak compactness (Theorems 5.23 and 5.24) and according to the w-S-continuity and w-uniform continuity arguments with features of this new nonstandard construction, we prove that there exist for every utility profile some $\mathcal{D}$-socially acceptable and σ-representative utility functions (Theorems 5.36 and 5.37).

Typically in the aggregation problems (which are used often in this work), we use some ultrafilters on the set of individuals to model the set of decisive coalitions. The more general case is suited to cater for those situations in aggregation theory in which some voters might abstain from voting, thus giving rise to profiles in which some coordinates might be empty[8]. Hence, there is one question:

> *Is the vote abstention situation satisfactorily covered by the generalised Kirman-Sondermann correspondence in [32]?*

At the moment, the vote abstention situation is not satisfactorily covered by the generalised Kirman-Sondermann correspondence in Herzberg and Eckert [32], because it is not reasonable to assume that the abstention of any single voter would force the outcome of an Arrovian rational aggregation to be empty. Therefore, for covering this gap, we apply a generalised definition of the ultraproduct in Makkai [41] (see Section 6.2) and generalise the Kirman-Sondermann correspondence by using this new type of ultraproduct in which empty models in some coordinates do not necessarily make the generalised ultraproduct also empty (Theorem 6.19). More technically, Makkai's ultraproduct yields the empty model unless non-empty models occur in each coordinate belonging to some member of its associated ultrafilter. Using this definition, we can give a more intuitive explanation for the situation of vote abstention.

[8]There are several ways to study and consider abstention in the social choice theory. For example, one can ignore any voters that abstain (see Pivato [50]), one can study abstention as if the voters ranked all candidates equally, and one can study abstention as if the voters submit two kind of inputs: ranking of candidates or abstention. In this work, we study the last approach.

Summary of contents

The plan of this thesis is as follows. In Chapter 2 we introduce the mathematical methods and techniques that are used in this work. We start with topics in mathematical logic and give some important results on ultrafilters (Ultrafilter Existence Theorem) and ultraproducts (Łoś Theorem). To accommodate the reader, we continue with a simple ultrapower construction of a nonstandard model of real numbers associated with a very understandable system of logic and prove the Basic Transfer Principle. Afterwards, we introduce the full theoretical background (superstructure and bounded ultrapowers) needed to establish nonstandard universe for modern developments and applications, like when we are dealing with mathematical objects such as Banach spaces and topological spaces. We also prove the important principles in this nonstandard universe (Transfer Principle, Internal Definition Principle, Overspill Principle and etc.).

Chapter 3 applies ultraproducts to find microfoundations for constructing a representative agent model. We construct a representative utility function for infinite-dimensional social decision problems, based on a bounded ultrapower construction with respect to the ultrafilter induced by the underlying social choice function.

In Chapter 4 we provide sufficient conditions for the preceding theorems in Chapter 3 to be satisfied in economic applications. After this modification, we show a possible macroeconomic application as an example.

Chapter 5 brings us to the theory of existence of a representative agent model for weaker conditions. We establish weaker conditions for our previous results and introduce an easy nonstandard approach to weak topological spaces. Then we prove the existence of a representative utility function with respect to this new constructed weak topology.

In Chapter 6 we study an aggregation problem with the model theoretic approach by means of generalised ultraproducts. We give preliminaries for Arrow-rational aggregators and generalised ultraproduct. Then we prove the generalised Kirman-Sondermann correspondence which covers the gap in Herzberg and Eckert's model-theoretic approach for voting abstention.

Finally, in Chapter 7 we give the conclusions of this thesis in a short and clear manner.

Chapter 2

Mathematical methodology

In this chapter we introduce the mathematical methods and techniques that are used in this work. We follow the notations and concepts from Bell and Slomson [10], Marker [46], Lauwers and Van Liedekerke [36], Albeverio et al. [2], Robinson [52], Goldblatt [24], and Wolff and Loeb [64].

2.1 Boolean algebra

One of the important concepts which will be used later to derive some consequences in predicate logic is the Boolean algebra. Moreover, the ultraproduct construction which has a main role in this work, is constructed by ultrafilters in power set Boolean algebras. Hence, for the convenience of the reader, we recall some notions on this topic.

2.1.1 Lattices

We have the following definition:

Definition 2.1. Let X be a non-empty set and $\leq$ be a binary relation on X. $\langle X, \leq \rangle$ is called a *partially ordered set* if and only if the following conditions are satisfied:

(1) Reflexivity: For all $x \in X$, $x \leq x$.

(2) Antisymmetric: For all $x, y \in X$, if $x \leq y$ and $y \leq x$, then $x = y$.

(3) Transitivity: For all $x, y, z \in X$, if $x \leq y$ and $y \leq z$, then $x \leq z$.

A partially ordered set $\langle X, \leq \rangle$ is called *totally ordered* if in addition to reflexivity, antisymmetric and transitivity, $\leq$ has the following property:

(4) Dichotomy: For all $x, y \in X$, either $x \leq y$ or $y \leq x$.

Definition 2.2. A subset A of a partially ordered set X is called a *chain* if and only if $\leq$ is dichotomy on A.

Example 2.1. Examples of partially and totally ordered sets are:

- Let X be any set. We denote the *power set* of X, i.e. the set of all subsets of X, by $\mathcal{P}(X)$. If $\subseteq$ is the relation of set inclusion, then $\langle \mathcal{P}(X), \subseteq \rangle$ is a partially ordered set.

- As an example of a totally ordered set, let $\mathbb{N} = \{0, 1, 2, \ldots\}$ be the set of natural numbers and let $\leq$ be the usual ordering of the natural numbers. Then $\langle \mathbb{N}, \leq \rangle$ is a totally ordered set.

For brevity, we denote the partially ordered set $\langle X, \leq \rangle$ only by X, when there is no risk of confusion.

Definition 2.3. If A is a subset of the partially ordered set X, an *upper bound* for A in X is an element of X greater than all the elements of A. The set of upper bounds of A is

$$\{u \in X \,:\, \forall a \in A \quad a \leq u\}.$$

Similarly a *lower bound* for A in X is an element of X less than all the elements of A. The set of lower bounds of A is

$$\{l \in X \,:\, \forall a \in A \quad l \leq a\}.$$

Definition 2.4. The *supremum* (or *least upper bound*) of A is an element $x \in X$, where x is an upper bound for A in X and is a lower bound for the set of all upper bounds for A in X. In other words, $x \in X$ is the supremum of A iff for all $a \in A$, $a \leq x$ and given any other upperbound y for A in X, we have $x \leq y$. We denote the supremum of A by $\sup(A)$.
Similarly $x \in X$ is called the *infimum* (or *greatest lower bound*) of A if it is a lower bound for A and greater than any other lower bound for A in X. We denote the infimum of A by $\inf(A)$.

A lattice is a partially ordered set in which the infimum and supremum of any two element set is defined, in other words:

Definition 2.5. A partially ordered set X is a *lattice* if and only if for all $x, y \in X$, $x \wedge y := \inf\{x, y\} \neq \emptyset$ and $x \vee y := \sup\{x, y\} \neq \emptyset$. $x \wedge y$ is called the *mee*t of x and y, and $x \vee y$ is called the *join* of x and y.

Example 2.2. Examples of lattices are:

- The partially ordered set $\langle \mathcal{P}(X), \subseteq \rangle$ of Example 2.1 is a lattice for any set X, since for any $A, B \in \mathcal{P}(X)$, $\sup\{A, B\} = A \cup B$, and $\inf\{A, B\} = A \cap B$.

- The space of all continuous real valued functions defined on a topological space X is a lattice, where the order relation for every two continuous real valued functions

f and g is defined by

$$f \leq g \quad \text{iff} \quad \forall x \in X \quad f(x) \leq g(x).$$

In this case, $\inf\{f,g\}(x) = \min\{f(x), g(x)\}$, and $\sup\{f,g\}(x) = \max\{f(x), g(x)\}$ for all $x \in X$.

A partially ordered set can only have one lower bound and one upper bound. Since a lattice is a partially ordered set, this is true also for the lattice and we call the unique upper bound of a lattice, the *maximum* of the lattice and denote it by 1. Similarly, we call the unique lower bound of a lattice, the *minimum* of the lattice and denote it by 0.

Definition 2.6. An element x of a lattice is said to be *maximal* if there is no y in the lattice, which is strictly greater than x.

Definition 2.7. For each $x \in L$, where L is a lattice, we define

$$\mathit{complement}(x) := \{\, y \in L \quad : \quad x \wedge y = 0, \quad x \vee y = 1 \,\},$$

as the set of complements of x. A lattice L is called *complemented* if and only if

(1) It has a maximum element 1 and a minimum element 0,

(2) For each $x \in L$, the set of complements of x is non-empty.

Remark 2.8. If the lattice L is complemented, then for all $x \in L$ we have:

(1) $x \wedge 0 = 0$,

(2) $x \vee 1 = 1$,

(3) $x \wedge 1 = x$,

(4) $x \vee 0 = x$.

We impose an extra condition on a complemented lattice to ensure the uniqueness of complements. In other words, a lattice can have more than one complement, except in the case that the meet and join operations are distributive:

Definition 2.9. A lattice L is said to be *distributive* if and only if

$$\forall x, y, z \in L \qquad (x \wedge y) \vee z = (x \vee z) \wedge (y \vee z),$$

and

$$\forall x, y, z \in L \qquad (x \vee y) \wedge z = (x \wedge z) \vee (y \wedge z).$$

Notation 2.10. *If a lattice is distributive and complemented, then the unique complement of an element x is denoted by x^*.*

Lemma 2.11. *If L is complemented and distributive, then each element has a unique complement.*

Proof. Suppose that $x \in L$ and since L is complemented, $complement(x)$ is non-empty and has at least one element. Now suppose that $y, z \in complement(x)$, then $x \vee y = 1$ and $x \wedge z = 0$. By using Remark 2.8, we have:

$$\begin{aligned} y &= y \vee 0 \\ &= y \vee (x \wedge z) && \text{(since } x \wedge z = 0) \\ &= (y \vee x) \wedge (y \vee z) && \text{(since L is distributive)} \\ &= 1 \wedge (y \vee z) && \text{(since } x \vee y = 1) \\ &= y \vee z. && \text{(by Remark 2.8)} \end{aligned}$$

Similarly one can show that $z = y \vee z$, and hence $y = z$. This completes the proof of the lemma. □

2.1.2 Boolean algebras

We have introduced above the complementation and distributivity of the lattices. Now we give the definition of the Boolean algebra according to this notions.

Definition 2.12. A *Boolean algebra* is a complemented and distributive lattice with at least two elements.

In other words, we can define a Boolean algebra $\mathcal{B}$ to be a structure

$$\mathcal{B} = \langle B, \vee, \wedge, ^*, 0, 1 \rangle,$$

where B is a set that contains the elements 0 and 1, and $\vee$, $\wedge$ and * are defined as above[1], satisfying the following conditions for every $x, y, z \in B$:

(1) (Associativity and commutativity of $\vee$ and $\wedge$):

$$x \vee (y \vee z) = (x \vee y) \vee z, \qquad x \wedge (y \wedge z) = (x \wedge y) \wedge z$$

and

$$x \vee y = y \vee x, \qquad x \wedge y = y \wedge x.$$

[1]We assume that in any Boolean algebra, 0 is different from 1.

(2) (Distributivity):

$$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z) \qquad \text{and} \qquad x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

(3) (Idempotent and absorbtion laws):

$$x \vee x = x, \qquad x \wedge x = x$$

and

$$x \vee (x \wedge y) = x, \qquad x \wedge (x \vee y) = x.$$

(4) (Laws of zero and one):

$$x \vee 0 = x, \qquad x \wedge 0 = 0, \qquad x \wedge x^* = 0$$

and

$$x \vee 1 = 1, \qquad x \wedge 1 = x, \qquad x \vee x^* = 1.$$

(5) (De Morgan laws and the law of double negation):

$$(x \vee y)^* = x^* \wedge y^*, \qquad (x \wedge y)^* = x^* \vee y^*$$

and

$$(x^*)^* = x.$$

If we had introduced Boolean algebra as above, we could define the relation $\leq$ on B by

$$x \leq y \quad \text{iff} \quad x \vee y = y.$$

Example 2.3. Examples of Boolean algebras are:

- The lattice of truth values $2 = \{0, 1\}$, with the partial ordering $\leq$ defined by $0 \leq 0$, $0 \leq 1$ and $1 \leq 1$ is a Boolean algebra. In other words, this is the smallest Boolean algebra $2 = \langle \{0, 1\}, \vee, \wedge, ^*, 0, 1 \rangle$ consists of only two elements.

- The partially ordered set $\langle \mathcal{P}(X), \subseteq \rangle$ in Example 2.2 is a Boolean algebra. In this case the maximum element is the set X and the minimum element is the empty set $\emptyset$, and complementation corresponds to set theoretic complementation, i.e. for any $C \in X, \bar{C} = X - C$. We call this Boolean algebra, a *power set Boolean algebra*. In other words, the power set Boolean algebra can be denoted as follows:

$$\mathcal{P}(X) = \langle 2^X, \cup, \cap, ^-, \emptyset, X \rangle,$$

where 2^X is set of all subsets of X, $\cup$ the union, $\cap$ the intersection and $^-$ the complement.

2.1.3 Filters and ultrafilters

Let L be a lattice and B a Boolean algebra.

Definition 2.13. A non-empty subset F of L is called a *filter* if and only if

(1) $x \wedge y \in F$ for all $x, y \in F$, and

(2) $z \in F$ whenever $x \leq z$ for all $x \in F$ and $z \in L$.

A filter F is called *proper* if and only if $F \neq L$.

Definition 2.14. A subset U of B is called an *ultrafilter* if and only if U is a maximal proper filter.

In other words, an ultrafilter U is a proper filter such that the only filter F with $U \subsetneq F$ is the whole Boolean algebra.

Definition 2.15. For every $A \subseteq B$, we define

- $A^0 := \{x \in B \,:\, \exists a \in A \quad a \leq x\}$
- $A^c := \{\inf(X) \,:\, X \subseteq A \quad X\, finite\}$.

If F is a filter, then a *base* of F is a set A with $A^0 = F$, and a *sub-base* of F is a set A such that A^c is a base for F, that is $A^{c0} := (A^c)^0 = F$. If A is a sub-base of F, then we also say that A generates F.

Remark 2.16. (1) $A \subseteq A^c$ and $A \subseteq A^0$ for all $A \subseteq B$, hence

$$A \subseteq A^c \subseteq A^{c0},$$

for all $A \subseteq B$. Also we have

$$A^{c0} = \{x \in B \,:\, \exists X \subseteq A \quad (X\, finite, \quad \inf(X) \leq x)\},$$

for every $A \subseteq B$.

(2) A filter F is proper if and only if $0 \notin F$.

(3) By definition, filters are closed under infima of pair-subsets.

Definition 2.17. Let $A \subseteq B$. We say that A has the *finite intersection property (fip)* if and only if $\inf(X) \neq 0$ for all finite $X \subseteq A$.

Then we have the following lemma:

Lemma 2.18. *Let $A \subseteq B$. Then*

(1) *A^{c0} is a filter.*

(2) *Any filter F containing A, contains A^{c0}.*

(3) *A^{c0} is a proper filter if and only if A has the finite intersection property.*

Proof. (1) Let $x, y \in A^{c0}$ and let $z \in B$ such that $x \leq z$. Then there exist finite subsets X, Y of A such that

$$\inf(X) \leq x, \quad \inf(Y) \leq y.$$

Then, $X \cup Y$ also is a finite subset of A and one can show that

$$\inf(X \cup Y) = \inf(X) \wedge \inf(Y) \leq x \wedge y,$$

which yields that $x \wedge y \in A^{c0}$. Now by transitivity of $\leq$, one has $\inf(X) \leq z$, so z also is in A^{c0}. Therefore A^{c0} is a filter.

(2) Let $F \supseteq A$ be a filter, and let $x \in A^{c0}$. Then there must be a finite subset $X \subseteq A$ such that $\inf(X) \leq x$. Since $A \subseteq F$, we have $X \subseteq F$, and since F is a filter, $\inf(X) \in F$ by Remark 2.16. Again using the fact that F is a filter, we have $x \in F$. Thus $A^{c0} \subseteq F$.

(3) Since A^{c0} is a filter (according to the first part), we need to prove that $A^{c0} \neq B$ if and only if A has the fip. Suppose $A^{c0} \neq B$. If A does not have the fip for some $X \subseteq A$, then $\inf(A) = 0$. Therefore $0 \in A^{c0}$ and so A^{c0} is not proper and this a contradiction with $A^{c0} \neq B$.

Conversely, suppose that A has the fip. If $A^{c0} = B$, then $0 \in A^{c0}$, and therefore there is some finite $X \subseteq A$ such that $\inf(X) \leq 0$, hence $\inf(X) = 0$. This contradicts the fip.

□

The following theorem about the characterisation of ultrafilters is fundamental for applications in mathematical logic:

Proposition 2.19. *Let U be a proper filter. U is an ultrafilter if and only if for every $x \in B$, either $x \in U$ or $x^* \in U$.*

Proof. First, suppose that U is an ultrafilter and let $x \notin U$. Define $G := (U \cup \{x\})^{c0}$. By the previous lemma, G is a filter which contains U. On the other hand, $x \in U \cup \{x\} \subseteq G$. Since U is an ultrafilter, G is not proper. Therefore $U \cup \{x\}$ does not have the fip and for some finite $X \subseteq U$, $\inf(X) \wedge x = 0$. This means that $\inf(X) \leq x^*$. But $\inf(X)$ is in U and hence $x^* \in U$ (by the definition of filter).

Conversely, suppose that for every $x \in B$, either $x \in U$ or $x^* \in U$. Let G be a filter with $U \subsetneq G$, and let $x \in G \backslash U$. Since $x \notin U$, $x^* \in U \subseteq G$. This shows that G is not proper. Hence U is a maximal proper filter, i.e. and ultrafilter. □

Definition 2.20. Let I be a set and $B = \mathcal{P}(I)$ (the power set Boolean algebra). Then for every $i \in I$, the set

$$\{X \in \mathcal{P}(I) \,:\, i \in X\},$$

is an ultrafilter. We call this ultrafilter the *principal ultrafilter* generated by i. An ultrafilter which is not principal is called *non-principal.*

Clearly, an ultrafilter on I is principal as soon it contains a finite subset of I. Hence, if I itself is finite, all ultrafilters on I are principal.

Theorem 2.21. (Ultrafilter Existence Theorem). *Every proper filter in a Boolean algebra can be extended to an ultrafilter.*

Proof. Let $F \neq B$ be a filter, and define

$$\mathcal{F} := \{G \subsetneq B \,:\, G \supseteq F, \quad G \,\text{filter}\}.$$

Since $F \in \mathcal{F}$, $\mathcal{F}$ is not empty. $(\mathcal{F}, \subseteq)$ is a partially ordered set. We will prove that with respect to this ordering, every chain in $\mathcal{F}$ has an upper bound in $\mathcal{F}$.

Let $\{D_i\}_{i \in I}$ be a chain in $\mathcal{F}$ and define $D := \cup_{i \in I} D_i$. For all $i \in I$, $D_i \subseteq D$ and $D \supseteq F$. Also, $0 \notin D_i$ for all $i \in I$ (as otherwise $D_i = B$ for some $i \in I$), hence $0 \notin D$. We claim that D is a filter: Let $x, y \in D$ and let $z \in B$ with $x \leq z$. There exist $i, j \in I$ such that $x \in D_i$ and $y \in D_j$. Since $\{D_i\}_{i \in I}$ is a chain, either $D_i \subseteq D_j$ or $D_j \subseteq D_i$. We may assume that $D_i \subseteq D_j$. Then, $x, y \in D_j$, which is a filter, hence $x \wedge y, z \in D_j \subseteq D$. Thus, D is a proper filter containing F, so $D \in \mathcal{F}$, and D is an upper bound for $\{D_i\}_{i \in I}$.

Hence, every chain in $\mathcal{F}$ has an upper bound in $\mathcal{F}$. We deduce from Zorn's lemma that $\mathcal{F}$ contains a maximal element U. In other words there can be no proper filter which includes U but is not equal to U, so U is an ultrafilter and this is the required ultrafilter which is an extension of F. □

Definition 2.22. Let I be a set and $B = \mathcal{P}(I)$. A subset $X \subseteq I$ is called *co-finite* if and only if $I \backslash X$ is finite.

Remark 2.23. Let I be an infinite set. The filter F of co-finite subsets of I can be extended to a non-principal ultrafilter U.

Proof. Since I is infinite, F is proper. By the previous theorem (Ultrafilter Existence Theorem), it can be extended to an ultrafilter U. If U were principal, then there would be some $i \in I$ such that $\{i\} \in U$ whilst on the other hand, $I\backslash\{i\}$ being co-finite would be in $F \subseteq U$. This would mean that

$$0 = \emptyset = \{i\} \cap (I\backslash\{i\}) \in U,$$

hence $U = \mathcal{P}(I)$, so U would not be an ultrafilter and this is a contradiction. □

2.2 Predicate Logic

In this section we introduce the formal language of predicate calculus (PC). We use this framework in order to define the notion of an interpretation and of a true formula.

2.2.1 Syntax of PC

Definition 2.24. The *language* $\mathcal{L}$ of PC is the set of symbols

$$\{\dot{v}_n \,:\, n \in \mathbb{N}\} \cup \{\dot{P}_n \,:\, n \in \mathbb{N}\} \cup \{\dot{=}, \dot{\exists}, \dot{\wedge}, \dot{\neg}, \dot{(}, \dot{)}\},$$

where $\dot{v}_n$ is called an *individual variable* for every $n \in \mathbb{N}$, $\dot{P}_n$ is called a *predicate letter* for every $n \in \mathbb{N}$, $\dot{=}$ is called the *equality symbol*, $\dot{\exists}$ is called the *existential quantifier*, $\dot{\wedge}$ and $\dot{\neg}$ are called *logical connectives*, $\dot{\exists}$, $\dot{\wedge}$ and $\dot{\neg}$ are called *logical symbols*, and $\dot{(}$ and $\dot{)}$ are called *punctuation symbols*.

Definition 2.25. A *string* of $\mathcal{L}$ is a finite sequence of symbols of $\mathcal{L}$. A string is an *atomic formula* if and only if it is of the form $\dot{v}_m \doteq \dot{v}_n$ or $\dot{P}_n\dot{(}\dot{v}_{l_0}, \ldots, \dot{v}_{l_{n-1}}\dot{)}$ where $m, n, l_0, \ldots, l_{n-1} \in \mathbb{N}$.

Definition 2.26. The set F of *formulae of PC* is the smallest subset of strings for which every atomic formula is in F, and for all $\phi, \psi \in F$ and every $n \in \mathbb{N}$, the strings

$$\dot{(}\dot{\exists}\dot{v}_n\dot{)}\phi, \quad \dot{(}\phi\dot{\wedge}\psi\dot{)}, \quad \dot{\neg}\phi,$$

are all in F.

Remark 2.27. For all $\phi, \psi \in F$ and every $n \in \mathbb{N}$, we introduce the following abbreviations:

$$\begin{aligned}
\dot{(}\dot{\forall}\dot{v}_n\dot{)}\phi &:= \dot{\neg}\dot{(}\dot{\exists}\dot{v}_n\dot{)}\dot{\neg}\phi; \\
\dot{(}\phi\dot{\vee}\psi\dot{)} &:= \dot{\neg}\dot{(}\dot{\neg}\phi\dot{\wedge}\dot{\neg}\psi\dot{)}; \\
\dot{(}\phi\dot{\rightarrow}\psi\dot{)} &:= \dot{\neg}\dot{(}\phi\dot{\wedge}\dot{\neg}\psi\dot{)}; \\
\dot{(}\phi\dot{\leftrightarrow}\psi\dot{)} &:= \dot{(}\dot{(}\phi\dot{\rightarrow}\psi\dot{)}\dot{\wedge}\dot{(}\psi\dot{\rightarrow}\phi\dot{)}\dot{)}.
\end{aligned}$$

When we are talking about the specific occurrence of some substring of a given string (e.g. a formula), we underline its position in the string. Let $m, n \in \mathbb{N}$ with $m \neq n$, and let ϕ, ψ be formulae. Moreover, suppose that X, Y are strings, and $\dot{v}_m$ is an individual variable. The *scope* of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in ϕ is defined as follows:

- The scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in $\underline{\dot{(}\dot{\exists}\dot{v}_n\dot{)}}\phi$ equals the formula $\dot{(}\dot{\exists}\dot{v}_n\dot{)}\phi$.
- The scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in $\dot{\neg}X\underline{\dot{(}\dot{\exists}\dot{v}_n\dot{)}}Y$ equals the scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in $X\underline{\dot{(}\dot{\exists}\dot{v}_n\dot{)}}Y$.
- The scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in $\dot{(}X\underline{\dot{(}\dot{\exists}\dot{v}_n\dot{)}}Y\dot{\wedge}\psi\dot{)}$ equals the scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in $X\underline{\dot{(}\dot{\exists}\dot{v}_n\dot{)}}Y$.
- The scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in $\dot{(}\dot{\exists}\dot{v}_m\dot{)}X\underline{\dot{(}\dot{\exists}\dot{v}_n\dot{)}}Y$ equals the scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in $X\underline{\dot{(}\dot{\exists}\dot{v}_n\dot{)}}Y$.

Definition 2.28. The occurrence of $\dot{v}_n$ in ϕ is said to be *bound* if and only if $\dot{v}_n$ occurs in the scope of some occurrence of $\dot{(}\dot{\exists}\dot{v}_n\dot{)}$ in ϕ. Otherwise, that occurrence of $\dot{v}_n$ is said to be *free* and we often write $\phi(\dot{v}_n)$.

Definition 2.29. A formula ϕ is called a *sentence* of $\mathcal{L}$ if and only if all occurrences of individual variables in ϕ are bound.

2.2.2 Semantics of PC

Definition 2.30. A *relational structure* is a pair $\mathfrak{A} = \langle A, (R_n)_{n\in\mathbb{N}}\rangle$, where A is a non-empty set, called the *domain* of $\mathfrak{A}$, and $R_n \subseteq A^n$ for every $n \in \mathbb{N}$. $\mathfrak{A}$ is called *$\mathcal{L}$-structure*[2].

Definition 2.31. If $A' \subseteq A$, then the *restriction* of $\mathfrak{A}$ to A' is the relational structure $\mathfrak{A}' = \langle A', (R'_n)_{n\in\mathbb{N}}\rangle$, where $R'_n = R_n \cap (A')^n$ for all $n \in \mathbb{N}$ and is denoted by $\mathrm{res}_{A'}\mathfrak{A}$.

Definition 2.32. If $\mathfrak{A}$ is an $\mathcal{L}$-structure, then any sequence $J : \mathbb{N} \to A$ is called a *valuation* on $\mathfrak{A}$. If J is a valuation, $m \in \mathbb{N}$ and $a \in A$, then we define $J(m/a)$ as follows:

$$J(m/a) : n \mapsto \begin{cases} J(n), & n \neq m \\ a, & n = m \end{cases}.$$

Now we have the following Tarski's definition of truth:

Definition 2.33. (Tarski's definition of truth). Let $\mathfrak{A}$ be an $\mathcal{L}$-structure, and let J be a valuation on $\mathfrak{A}$. We define the predicate (or relation) $\mathfrak{A} \models_J \phi$, which is read as 'J satisfies ϕ in $\mathfrak{A}$', for all $\mathcal{L}$-formulae ϕ, ψ and $m, n, l_0, \dots, l_{n-1} \in \mathbb{N}$, as follows:

[2] $\mathfrak{A}$ is also called a realisation of $\mathcal{L}$.

- $\mathfrak{A} \models_J \dot{v}_m \doteq \dot{v}_n$ if and only if $J(m) = J(n)$,
- $\mathfrak{A} \models_J \dot{P}_n \dot{(} \dot{v}_{l_0}, \ldots, \dot{v}_{l_{n-1}} \dot{)}$ if and only if $(J(l_0), \ldots, J(l_{n-1})) \in R_n$,
- $\mathfrak{A} \models_J \dot{(} \dot{\exists} \dot{v}_n \dot{)} \phi$ if and only if there exists some $a \in A$ such that $\mathfrak{A} \models_{J(n/a)} \phi$,
- $\mathfrak{A} \models_J \dot{\neg} \phi$ if and only if $\mathfrak{A} \not\models_J \phi$[3],
- $\mathfrak{A} \models_J \phi \dot{\wedge} \psi$ if and only if $\mathfrak{A} \models_J \phi$ and $\mathfrak{A} \models_J \psi$.

Definition 2.34. Let λ be a sentence and let Λ be a set of sentences. $\mathfrak{A}$ is a *model* of λ if and only if $\mathfrak{A} \models_J \lambda$ for some (and hence every) valuation J. In this case we write $\mathfrak{A} \models \lambda$ and we say that λ is *true* or *valid* in $\mathfrak{A}$.
$\mathfrak{A}$ is a model of Λ if and only if there exists some J such that $\mathfrak{A} \models_J \lambda$ for every $\lambda \in \Lambda$. In this case we write $\mathfrak{A} \models \Lambda$.

Definition 2.35. Let λ be a sentence. λ is *universally valid* if $\mathfrak{B} \models \lambda$ for all $\mathcal{L}$-structures $\mathfrak{B}$.

Notation 2.36. *Let $\mathfrak{A}$ be an $\mathcal{L}$-structure, and let J be a valuation on $\mathfrak{A}$. Suppose ϕ is a formula with n free variables $\dot{v}_0, \ldots, \dot{v}_{n-1}$ such that the valuation J assigns $\dot{v}_j$ the value $x_j \in A$ for all $j \leq n-1$. We shall often write*

$$\mathfrak{A} \models \phi[x_0, \ldots, x_{n-1}],$$

instead of $\mathfrak{A} \models_J \phi$, and say $\phi[x_0, \ldots, x_{n-1}]$ is true in $\mathfrak{A}$.

2.3 Ultraproducts

In this section we introduce the ultraproduct construction which will be of high relevance in the rest of this work. Let I be a set. A filter and an ultrafilter on I is respectively a filter and an ultrafilter in the power set Boolean algebra $\mathcal{P}(I)$. We fix some filter F on I.

We fix a relational structure $\mathfrak{A}^i = \langle A^i, (R^i_n)_{n \in \mathbb{N}} \rangle$ for every $i \in I$, which is adapted to the language $\mathcal{L}$ of PC. We write $\prod_{i \in I} A^i$ for the cartesian product of all A^i. Where the index set is clear from the context, we write just $\prod A^i$. If $\underline{a} \in \prod A^i$, we denote its i-th coordinate by $\underline{a}(i)$.

2.3.1 Reduced products and ultraproducts

At first we define the *binary relation* $\sim_F$ on $\prod_{i \in I} A^i$ as follows:

[3] $\mathfrak{A} \not\models_J \phi$ is read as 'J does not satisfy ϕ in $\mathfrak{A}$'.

Definition 2.37. For every $\underline{a}, \underline{b} \in \prod_{i\in I} A^i$ we have

$$\underline{a} \sim_F \underline{b} \Leftrightarrow \{i \in I \,:\, \underline{a}(i) = \underline{b}(i)\} \in F.$$

Remark 2.38. $\sim_F$ is an equivalence relation.

Proof. Since F is a filter on I, $I \in F$ and since $=$ is symmetric and reflexive, the relation $\sim_F$ is also symmetric and reflexive. We will show that it is also a transitive relation. Suppose that $\underline{a} \sim_F \underline{b}$ and $\underline{b} \sim_F \underline{c}$ for all $\underline{a}, \underline{b}, \underline{c} \in \prod_{i\in I} A^i$. Then by definition of $\sim_F$ we have

$$\{i \in I \,:\, \underline{a}(i) = \underline{b}(i)\} \in F,$$

and

$$\{i \in I \,:\, \underline{b}(i) = \underline{c}(i)\} \in F.$$

Since F is a filter, it is closed under intersections. Hence

$$\{i \in I \,:\, \underline{a}(i) = \underline{b}(i)\} \cap \{i \in I \,:\, \underline{b}(i) = \underline{c}(i)\} \in F.$$

But

$$\begin{aligned} F \ni \{i \in I \,:\, \underline{a}(i) = \underline{b}(i)\} \cap \{i \in I \,:\, \underline{b}(i) = \underline{c}(i)\} &= \{i \in I \,:\, \underline{a}(i) = \underline{b}(i) = \underline{c}(i)\} \\ &\subseteq \{i \in I \,:\, \underline{a}(i) = \underline{c}(i)\}. \end{aligned}$$

Since F is closed under supersets, $\{i \in I \,:\, \underline{a}(i) = \underline{c}(i)\} \in F$, which means $\underline{a} \sim_F \underline{c}$. Thus, $\sim_F$ is also transitive. □

Now we define a relation S on $\prod_{i\in I} A^i$ as follows:

Definition 2.39. For every $\underline{a}, \underline{b} \in \prod_{i\in I} A^i$,

$$(\underline{a}, \underline{b}) \in S \Leftrightarrow \{i \in I \,:\, (\underline{a}(i), \underline{b}(i)) \in R^i\} \in F.$$

Definition 2.40. For all $n \in \mathbb{N}$, define

$$S_n := \left\{ \left(\underline{a}_0, \ldots, \underline{a}_{n-1}\right) \,:\, \{i \in I \,:\, \left(\underline{a}_0(i), \ldots, \underline{a}_{n-1}(i)\right) \in R^i_n\} \in F \right\}.$$

Lemma 2.41. *For all $n \in \mathbb{N}$ and for all $\underline{a}_0, \ldots, \underline{a}_{n-1}, \underline{b}_0, \ldots, \underline{b}_{n-1} \in \prod_{i\in I} A^i$, if $(\underline{a}_0, \ldots, \underline{a}_{n-1}) \in S_n$ and $\underline{a}_l \sim_F \underline{b}_l$ for all $l \leq n-1$, then $(\underline{b}_0, \ldots, \underline{b}_{n-1}) \in S_n$. It means that $\sim_F$ is a* congruence relation.

Proof. For $n \in \mathbb{N}$ and $\underline{a}_0, \ldots, \underline{a}_{n-1}, \underline{b}_0, \ldots, \underline{b}_{n-1} \in \prod_{i\in I} A^i$, we have

$$\{i \in I \,:\, (\underline{a}_0(i), \ldots, \underline{a}_{n-1}(i)) \in R^i_n\} \cap \bigcap_{l=0}^{n-1} \{i \in I \,:\, \underline{a}_l(i) = \underline{b}_l(i)\}.$$

Since $(\underline{a}_0, \ldots, \underline{a}_{n-1}) \in S_n$, $\{i \in I : \big(\underline{a}_0(i), \ldots, \underline{a}_{n-1}(i)\big) \in R_n^i\}$ is in F, and by definition of $\sim_F$, $\{i \in I : \underline{a}_l(i) = \underline{b}_l(i)\}$ is also in F. But F, as a filter on I, is closed under intersections, therefore

$$\{i \in I : (\underline{a}_0(i), \ldots, \underline{a}_{n-1}(i)) \in R_n^i\} \cap \bigcap_{l=0}^{n-1} \{i \in I : \underline{a}_l(i) = \underline{b}_l(i)\} \in F.$$

The last intersection can be written as follows:

$$\{i \in I : \underline{a}_0(i) = \underline{b}_0(i), \ldots, \underline{a}_{n-1}(i) = \underline{b}_{n-1}(i), \quad (\underline{a}_0(i), \ldots, \underline{a}_{n-1}(i)) \in R_n^i\},$$

which is a subset of

$$\{i \in I : (\underline{b}_0(i), \ldots, \underline{b}_{n-1}(i)) \in R_n^i)\}.$$

The filter F is also closed under supersets, this means that

$$\{i \in I : (\underline{b}_0(i), \ldots, \underline{b}_{n-1}(i)) \in R_n^i)\} \in F,$$

and hence $(\underline{b}_0, \ldots, \underline{b}_{n-1}) \in S_n$. □

We have the following important definition:

Definition 2.42. For all $\underline{a} \in \prod A^i$, the *equivalence class of* $\underline{a}$ *with respect to* F is defined as follows:

$$[\underline{a}]_F := [\underline{a}]_{\sim_F} = \left\{\underline{b} \in \prod A^i : \underline{b} \sim_F \underline{a}\right\}.$$

We use the following notation for the collection of all equivalence classes of the elements of $\prod A^i$:

$$\prod A^i/F := \left\{[\underline{a}]_F : \underline{a} \in \prod A^i\right\}.$$

For all $n \in \mathbb{N}$, let

$$R_n^F := \{([\underline{a}_1]_F, \ldots, [\underline{a}_n]_F) : (\underline{a}_1, \ldots, \underline{a}_n) \in S_n)\}.$$

Now $\prod \mathfrak{A}^i/F := \langle \prod A^i/F, (R_n^F)_{n \in \mathbb{N}} \rangle$ is called the *reduced product* of the $\mathfrak{A}^i$ with respect to F. If the $\mathfrak{A}^i$ are all equal, then $\mathfrak{A}^I/F = \prod \mathfrak{A}^i/F$ is said to be a *reduced power.*

Note that, since $\sim_F$ is a congruence relation, R_n^F is well-defined for every $n \in \mathbb{N}$. Hence the reduced product $\prod \mathfrak{A}^i/F$ is an L-structure.

Definition 2.43. In the previous definition, if F is an ultrafilter, then $\prod \mathfrak{A}^i/F$ is called an *ultraproduct* and $\mathfrak{A}^I/F$ is said to be an *ultrapower.*

2.3.2 Łoś's Theorem

We consider the following definition for every sequence $\underline{\underline{x}} = (\underline{x}_0, \underline{x}_1, \ldots) \in \left(\prod A^i\right)^{\mathbb{N}}$:

$$[\underline{\underline{x}}]_F := ([\underline{x}_0]_F, [\underline{x}_1]_F, \ldots) \in \left(\prod A^i/F\right)^{\mathbb{N}}.$$

We let $\underline{\underline{x}}(i) = (\underline{x}_0(i), \underline{x}_1(i), \ldots)$ be the sequence of elements of A^i, that is $\underline{\underline{x}}(i) \in \left(A^i\right)^{\mathbb{N}}$.

The most important and fundamental result on ultraproducts is Łoś's Theorem which was established in 1955 [38] by the Polish mathematician, logician, economist and philosopher Jerzy Łoś.

Theorem 2.44. (Łoś's Theorem). *If F is an ultrafilter, then for every $\mathcal{L}$-formula ϕ and every sequence $\underline{\underline{x}} \in \left(\prod A^i\right)^{\mathbb{N}}$,*

$$\prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} \phi \Leftrightarrow \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \phi\} \in F.$$

Proof. The proof follows by induction with respect to the complexity of ϕ, i.e., the number of logical symbols in ϕ. Suppose ϕ is an atomic formula of the form $\phi = \dot{v}_m \doteq \dot{v}_n$ for some $m, n \in \mathbb{N}$, then by Tarski's definition of truth we have

$$\prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} \dot{v}_m \doteq \dot{v}_n \Leftrightarrow [\underline{x}_m]_F = [\underline{x}_n]_F.$$

But $[\underline{x}_m]_F = [\underline{x}_n]_F$ if and only if $\underline{x}_m \sim_F \underline{x}_n$, and by the definition of $\sim_F$, $\underline{x}_m \sim_F \underline{x}_n$ if and only if $\{i \in I : \underline{x}_m(i) = \underline{x}_n(i)\} \in F$. Again by Tarski's definition of truth, $\{i \in I : \underline{x}_m(i) = \underline{x}_n(i)\} \in F$ if and only if

$$\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \dot{v}_m \doteq \dot{v}_n\} \in F.$$

Therefore, for the atomic formula of the form $\phi = \dot{v}_m \doteq \dot{v}_n$, we have shown

$$\prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} \phi \Leftrightarrow \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \phi\} \in F.$$

Suppose that ϕ is an atomic formula of the form $\phi = \dot{P}_n \dot{(} \dot{v}_{l_0}, \ldots, \dot{v}_{l_{n-1}} \dot{)}$ for some $n \in \mathbb{N}$ and $l_0, \ldots, l_{n-1} \in \mathbb{N}$, then, using Tarski's definition of truth we have

$$\prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} \dot{P}_n \dot{(} \dot{v}_{l_0}, \ldots, \dot{v}_{l_{n-1}} \dot{)} \Leftrightarrow \left([\underline{x}_{l_0}]_F, \ldots, [\underline{x}_{l_{n-1}}]_F\right) \in R_n^F.$$

But $\left([\underline{x}_{l_0}]_F, \ldots, [\underline{x}_{l_{n-1}}]_F\right) \in R_n^F$ if and only if $(\underline{x}_{l_0}, \ldots, \underline{x}_{l_{n-1}}) \in S_n$, and by the definition of S_n, we have

$$(\underline{x}_{l_0}, \ldots, \underline{x}_{l_{n-1}}) \in S_n \Leftrightarrow \{i \in I : (\underline{x}_{l_0}(i), \ldots, \underline{x}_{l_{n-1}}(i)) \in R_n^i\} \in F.$$

Again, by Tarski's definition of truth, $\{i \in I : (\underline{x}_{l_0}(i), \ldots, \underline{x}_{l_{n-1}}(i)) \in R_n^i\} \in F$ if and only if $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \dot{P}_n\dot{(}\dot{v}_{l_0}, \ldots, \dot{v}_{l_{n-1}}\dot{)}\} \in F$. Hence, for the atomic formula of the form $\phi = \dot{P}_n\dot{(}\dot{v}_{l_0}, \ldots, \dot{v}_{l_{n-1}}\dot{)}$ we have shown:

$$\prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \phi \Leftrightarrow \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \phi\} \in F.$$

Therefore the result is true for all atomic formulae. This forms the basis of the induction. Now, we prove the induction step of the proof. Suppose that the result is true for all formulae containing less than r logical symbols and that ϕ is a formula containing r logical symbols. There are three cases to consider:

(1) If $\phi = \dot{\neg}\psi$ for some $\mathcal{L}$ formula ψ. Then by Tarski's definition of truth we have

$$\prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \dot{\neg}\psi \Leftrightarrow \prod \mathfrak{A}^i/F \not\models_{[\underline{x}]_F} \psi.$$

By using the induction hypothesis, $\prod \mathfrak{A}^i/F \not\models_{[\underline{x}]_F} \psi$ if and only if

$$\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi\} \notin F.$$

Since F is an ultrafilter,

$$\begin{aligned} \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi\} \notin F &\Leftrightarrow I \setminus \Big(\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi\}\Big) \in F \\ &\Leftrightarrow \{i \in I : \mathfrak{A}^i \not\models_{\underline{\underline{x}}(i)} \psi\} \in F. \end{aligned} \tag{2.1}$$

Again by Tarski's definition of truth, $\{i \in I : \mathfrak{A}^i \not\models_{\underline{\underline{x}}(i)} \psi\} \in F$ if and only if $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \dot{\neg}\psi\} \in F$. Hence for the formula of the form $\phi = \dot{\neg}\psi$, we have shown

$$\prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \phi \Leftrightarrow \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \phi\} \in F.$$

(2) If $\phi = \psi \dot{\wedge} \chi$ for some $\mathcal{L}$ formulae ψ and χ. Then by Tarski's definition of truth, we have

$$\prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \psi \dot{\wedge} \chi \Leftrightarrow \prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \psi \quad \text{and} \quad \prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \chi.$$

By using the induction hypothesis, $\prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \psi$ and $\prod \mathfrak{A}^i/F \models_{[\underline{x}]_F} \chi$ if and only if $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi\} \in F$ and $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \chi\} \in F$ respectively. Since F is closed under finite intersections,

$$\begin{aligned} &\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi\} \cap \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \chi\} \in F \\ \Leftrightarrow &\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi, \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \chi\} \in F. \end{aligned}$$

Again by Tarski's definition of truth, $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi, \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \chi\} \in F$ if and only if $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \psi \dot{\wedge} \chi\} \in F$. Hence for the formula of the form $\phi = \psi \dot{\wedge} \chi$, we have shown

$$\prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} \phi \Leftrightarrow \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \phi\} \in F.$$

(3) If $\phi = \dot{(}\dot{\exists}\dot{v}_n\dot{)}\psi$ for some $n \in \mathbb{N}$ and some $\mathcal{L}$ formula ψ. Then by using Tarski's definition of truth we get

$$\begin{aligned} \prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} \dot{(}\dot{\exists}\dot{v}_n\dot{)}\psi &\Leftrightarrow \exists [\underline{a}]_F \in \prod A^i/F \qquad \prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F(n/[\underline{a}]_F)} \psi \\ &\Leftrightarrow \exists \underline{a} \in \prod A^i \qquad \prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F(n/[\underline{a}]_F)} \psi \\ &\Leftrightarrow \exists \underline{a} \in \prod A^i \qquad \prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}(n/\underline{a})]_F} \psi. \end{aligned}$$

By using the induction hypothesis,

$$\exists \underline{a} \in \prod A^i \qquad \prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}(n/\underline{a})]_F} \psi,$$

if and only if

$$\exists \underline{a} \in \prod A^i \qquad \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F.$$

Since there exists an $\underline{a} \in \prod A^i$ with $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F$, clearly

$$\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \subseteq \{i \in I : \exists \underline{a}(i) \in A^i \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\},$$

and since $\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F$ and F is closed under supersets, we get $\{i \in I : \exists \underline{a}(i) \in A^i \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F$. On the other hand, if $\{i \in I : \exists \underline{a}(i) \in A^i \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F$, then by the Axiom of Choice, there exists some $\underline{a} \in \prod A^i$ such that

$$\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \supseteq \{i \in I : \exists \underline{a}(i) \in A^i \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F.$$

Since F is closed under supersets, we have found some $\underline{a} \in \prod A^i$ with

$$\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F.$$

Therefore we have shown that

$$\exists \underline{a} \in \prod A^i \qquad \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F$$

if and only if

$$\{i \in I : \exists \underline{a}(i) \in A^i \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F.$$

Hence

$$\prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} (\dot{\exists} \dot{v}_n \dot{)} \psi \Leftrightarrow \{i \in I : \exists \underline{a}(i) \in A^i \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F.$$

Now by using again Tarski's definition of truth,

$$\{i \in I : \exists \underline{a}(i) \in A^i \quad \mathfrak{A}^i \models_{\underline{\underline{x}}(i)(n/\underline{a}(i))} \psi\} \in F$$

if and only if

$$\{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} (\dot{\exists} \dot{v}_n \dot{)} \psi\} \in F.$$

Therefore for the formula of the form $\phi = (\dot{\exists} \dot{v}_n \dot{)} \psi$, we have shown

$$\prod \mathfrak{A}^i/F \models_{[\underline{\underline{x}}]_F} \phi \Leftrightarrow \{i \in I : \mathfrak{A}^i \models_{\underline{\underline{x}}(i)} \phi\} \in F,$$

and this completes the proof of the Łoś theorem.

□

2.4 An ultrapower construction of a nonstandard model of the real numbers

In this section we extend the real numbers $\mathbb{R}$ to the nonstandard real numbers ${}^*\mathbb{R}$ which contain infinitesimals and infinitely large numbers. At first we will do this in a simple way[4].

Recall from Section 2.1.3 that a *non-principal ultrafilter* on $\mathbb{N}$ is a collection $F \subset \mathcal{P}(\mathbb{N})$ such that

(1) $\emptyset \notin F$,

(2) If $F_1, F_2 \in F$, then $F_1 \cap F_2 \in F$,

(3) If $F_1 \subset \mathbb{N}$ and $F_1 \notin F$, then $\mathbb{N} \backslash F_1 \in F$,

(4) If F_1 is a finite subset of $\mathbb{N}$, then $\mathbb{N} \backslash F_1 \in F$.

From Properties (1), (2) and (3) one can deduce the property that any superset of a set in F is also in F, i.e., F is closed under supersets[5]. For proving this, suppose that

[4] We will study the general case in Section 2.6.

[5] That is, if $F_1 \subseteq F_2$ and $F_1 \in F$, then $F_2 \in F$.

$F_2 \notin F$. Hence by Property (3), $\mathbb{N} \backslash F_2 \in F$. By using Property (2), $\emptyset = F_1 \cap \mathbb{N} \backslash F_2 \in F$, which is a contradiction with Property (1).

If we replace Property (3) with this weak property (F is closed under supersets), then, given Property (4), we have just a *free filter* on $\mathbb{N}$. In other words, an ultrafilter on $\mathbb{N}$ is the collection of all subsets of $\mathbb{N}$ which have ν-measure 1 for some ν, which is a finitely additive $\{0,1\}$-valued measure on $2^{\mathbb{N}}$.

Let F be a non-principal ultrafilter on $\mathbb{N}$ and we have the following equivalence relation for sequences x and y in $\mathbb{R}^{\mathbb{N}}$:

$$x \sim_F y \quad \text{iff} \quad \{i \in \mathbb{N} : x(i) = y(i)\} \in F. \tag{2.2}$$

Definition 2.45. Given $x \in \mathbb{R}^{\mathbb{N}}$, let $[x]_F$ denote the equivalence class of x with respect to the equivalence relation $\sim_F$. The set of *nonstandard real numbers* or *hyperreals*, denoted by $^*\mathbb{R}$, is the collection of such equivalence classes. In symbols,

$$^*\mathbb{R} = \mathbb{R}^{\mathbb{N}}/F = \{[x]_F : x \in \mathbb{R}^{\mathbb{N}}\}.$$

If $x \in \mathbb{R}^{\mathbb{N}}$, we denote its image in $^*\mathbb{R}$ by $[x]_F$. Of course, every element in $^*\mathbb{R}$ is of the form $[x]_F$ for some $x \in \mathbb{R}^{\mathbb{N}}$.

For any real number $r \in \mathbb{R}$, let $\mathbf{r} = (r, r, \ldots)$ be the constant sequence in $\mathbb{R}^{\mathbb{N}}$. Then we have an embedding

$$^* : \mathbb{R} \to {}^*\mathbb{R}, \tag{2.3}$$

where $^*r = [\mathbf{r}]_F$ for every $r \in \mathbb{R}$.

Remark 2.46. In the last definition, $\mathbb{R}^{\mathbb{N}}/F$ is an ultrapower construction which is defined in Section 2.3.

The structure $\langle \mathbb{R}, +, \times, <, 0, 1 \rangle$ is an ordered field, where $\mathbb{R}$ is the set of elements of the structure, $+$ and $\times$ are the binary operations of addition and multiplication, $<$ is the ordering relation, and 0 and 1 are two distinguished elements of the domain.

By Embedding (2.3), we have $^*0 = [\mathbf{0}]_F$ and $^*1 = [\mathbf{1}]_F$. From the Equivalence (2.2), we can introduce the equality of any two elements $[x]_F, [y]_F$ in $^*\mathbb{R}$ as follows:

$$[x]_F \; ^*= \; [y]_F \qquad \text{iff} \qquad \{i \in \mathbb{N} : x(i) = y(i)\} \in F. \tag{2.4}$$

In a similar way we extend the ordering relation $<$ to $^*\mathbb{R}$ by setting for every $[x]_F, [y]_F \in {}^*\mathbb{R}$:

$$[x]_F \; ^*< \; [y]_F \qquad \text{iff} \qquad \{i \in \mathbb{N} : x(i) < y(i)\} \in F. \tag{2.5}$$

Remark 2.47. The relation $^*<$ on $^*\mathbb{R}$ extends the relation $<$ on $\mathbb{R}$, i.e., given any $x, y \in \mathbb{R}$ we see that $x < y$ in $\mathbb{R}$ if and only if $^*x \; ^*< \; ^*y$ in $^*\mathbb{R}$.

In addition to $=$ and $<$, we extend the operations $+$ and $\cdot$ for any two elements $[x]_F, [y]_F$ in ${}^*\mathbb{R}$ in a similar way:

$$[x]_F \ {}^*+\ [y]_F \ {}^*=\ [z]_F \quad \text{iff} \quad \{i \in \mathbb{N} : x(i) + y(i) = z(i)\} \in F, \tag{2.6}$$
$$[x]_F \ {}^*\times\ [y]_F \ {}^*=\ [z]_F \quad \text{iff} \quad \{i \in \mathbb{N} : x(i) \times y(i) = z(i)\} \in F. \tag{2.7}$$

Now, we have the following proposition which will be proved later, after stating the Basic Transfer Principle.

Proposition 2.48. *The structure* $\langle {}^*\mathbb{R}, {}^*+, {}^*\times, {}^*<, {}^*0, {}^*1\rangle$ *is an ordered field extension of the ordered field* $\langle \mathbb{R}, +, \times, <, 0, 1\rangle$.

From the previous proposition, we can verify that ${}^*\mathbb{R}$ contains infinitesimals and infinite numbers.

Convention 2.49. Henceforth, for the ease of reading, common operation and relation symbols on ${}^*\mathbb{R}$, such as ${}^*+, {}^*\times, {}^*<, \ldots$, are simply written as $+, \times, <, \ldots$. We also denote the elements of ${}^*\mathbb{R}$ without the asterisk.

Definition 2.50. Let $\delta \in {}^*\mathbb{R}$. Then we say δ is a *(positive) infinitesimal* if ${}^*0 < \delta < {}^*r$ for all $r > 0$ in $\mathbb{R}$.

Let $f : \mathbb{R} \times \ldots \times \mathbb{R} \to \mathbb{R}$ be an n-ary function. We introduce as before the extended function ${}^*f$ by the equivalence:

$${}^*f\left([x_0]_F, \ldots, [x_{n-1}]_F\right) = [y]_F \quad \text{iff} \quad \{i \in \mathbb{N} : f(x_0(i), \ldots, x_{n-1}(i)) = y(i)\} \in F. \tag{2.8}$$

Remark 2.51. The function ${}^*f$ extends f, i.e., given any $x \in \mathbb{R}^n$, $f(x_0, \ldots, x_{n-1}) = x$ if and only if ${}^*f({}^*x_0, \ldots, {}^*x_{n-1}) = {}^*x$.

Due to Equivalence (2.8), we can extend the absolute value function to ${}^*\mathbb{R}$ and denote it by $|\cdot|$ instead of ${}^*|\cdot|$. We have the following important definition:

Definition 2.52. An element $x \in {}^*\mathbb{R}$ is called *finite* if $|x| < {}^*r$ for some $r > 0$ in $\mathbb{R}$. We let $\mathrm{Fin}({}^*\mathbb{R})$ denote the finite elements of ${}^*\mathbb{R}$.

Definition 2.53. Given $x \in {}^*\mathbb{R}$, we say that x is *infinitely close* to some $r \in \mathbb{R}$, in notation $x \approx r$, if $|x - {}^*r|$ is infinitesimal in ${}^*\mathbb{R}$.

There is an alternative definition for finite and infinitesimal elements in ${}^*\mathbb{R}$:

Definition 2.54. For any $x \in {}^*\mathbb{R}$,

- x is *finite* if $|x| < n$ for some $n \in \mathbb{N}$;
- x is *infinitesimal* if $|x| < \frac{1}{n}$ for every $n \in \mathbb{N} \setminus \{0\}$.

Note that 0 is the only infinitesimal in $\mathbb{R}$. Clearly, $\approx$ is an equivalence relation on ${}^*\mathbb{R}$, and we have the following definition:

Definition 2.55. Let $x \in {}^*\mathbb{R}$. The equivalence class of x with respect to the relation $\approx$ is called the *monad* of x and denoted by:

$$\mu(x) = \{y \in {}^*\mathbb{R} : y \approx x\}.$$

Proposition 2.56. *Let $x \in {}^*\mathbb{R}$ be finite, i.e., $x \in Fin({}^*\mathbb{R})$. Then there exists a unique $r \in \mathbb{R}$ such that $x \approx r$.*

Proof. Fix a finite $x \in {}^*\mathbb{R}$, and set $A := \{y \in \mathbb{R} : {}^*y \leq x\}$. Here we have used x to define an upper bounded subset of ordinary real numbers. Let r be the least upper bound in $\mathbb{R}$ of the set A. Assume $|{}^*r - x|$ is not infinitesimal. Then for some $n \in \mathbb{N}\backslash\{0\}$, $|{}^*r - x| \geq \frac{1}{n}$. In this case, if ${}^*r < x$, then $r + \frac{1}{n}$ is still in A, so r is not the least upper bound of A. On the other hand, if ${}^*r > x$, then $r - \frac{1}{n}$ is an upper bound of A, and this again contradicts the definition of r. It follows that, $r \approx x$.
Now let $r_1, r_2 \in \mathbb{R}$ and $r_1 \approx x \approx r_2$, then $r_1 \approx r_2$, so $|r_1 - r_2|$ is infinitesimal. Since 0 is the only infinitesimal in $\mathbb{R}$, $r_1 = r_2$. □

By the previous proposition, the following definition is well-defined:

Definition 2.57. If $x \in {}^*\mathbb{R}$ is finite, then the unique real number r with $x \approx r$ is called the *standard part* of x. We write ${}^\circ x = r$.

We have already noticed that $\mathbb{R}$ and ${}^*\mathbb{R}$ are similar in many respects, for example, both are ordered fields. We shall now make precise in which sense $\mathbb{R}$ and ${}^*\mathbb{R}$ are similar. We will specify in more detail which properties of $\mathbb{R}$ transfer to ${}^*\mathbb{R}$. For this, we need a formal language for an ordered field (OF).

The structure $\mathbb{R}$ has a language $\mathcal{L}_{OF}$ that can be used to describe the kind of properties of $\mathbb{R}$ that are preserved under the embedding ${}^* : \mathbb{R} \to {}^*\mathbb{R}$. The language $\mathcal{L}_{OF}$ is the set of symbols

$$\{\dot{x}, \dot{y}, \dot{z}, \dot{x}_0, \dot{x}_1, \ldots, \dot{y}_0, \dot{y}_1, \ldots\} \cup \{\dot{0}, \dot{1}, \dot{+}, \dot{\times}, \dot{<}\} \cup \{\dot{=}, \dot{\exists}, \dot{\wedge}, \dot{\neg}, \dot{(}, \dot{)}\},$$

where

- $\dot{x}, \dot{y}, \dot{z}, \dot{x}_0, \dot{x}_1, \ldots, \dot{y}_0, \dot{y}_1, \ldots$ are *variables*,
- $\dot{0}$ and $\dot{1}$ are *unary predicate letters*,
- $\dot{+}$ and $\dot{\times}$ are *binary operations*,
- $\dot{<}$ is a *binary relation*,

- $\dot{=}$ is the *equality symbol*,
- $\dot{\exists}$ is the *existential quantifier*,
- $\dot{\wedge}$ and $\dot{\neg}$ are *logical connectives*, and
- $\dot{(}$ and $\dot{)}$ are *parentheses.*

A string is an *atomic formula* if and only if it is of the form $\dot{x} \doteq \dot{y}$, $\dot{x} \dot{+} \dot{y}$, $\dot{x}\dot{\times}\dot{y}$ or $\dot{x}\dot{<}\dot{y}$. The set of *formulae* is the smallest subset of strings for which every atomic formula is a formula of $\mathcal{L}_{OF}$, and if ϕ, ψ are formulae of $\mathcal{L}_{OF}$ and $\dot{x}$ is a variable, then

$$\dot{\neg}\phi, \quad \dot{(}\phi\dot{\wedge}\psi\dot{)}, \quad \dot{(}\dot{\exists}\dot{x}\dot{)}\phi,$$

are formulae of $\mathcal{L}_{OF}$.

For all formulae ϕ, ψ of $\mathcal{L}_{OF}$ and all variables $\dot{x}$ and $\dot{y}$, we introduce the following abbreviations:

$$\begin{aligned}
\dot{(}\dot{x}\dot{\neq}\dot{y}\dot{)} &:= \dot{\neg}\dot{(}\dot{x} \doteq \dot{y}\dot{)};\\
\dot{(}\dot{x}\dot{>}\dot{y}\dot{)} &:= \dot{\neg}\dot{(}\dot{x}\dot{<}\dot{y}\dot{)};\\
\dot{(}\dot{\forall}\dot{x}\dot{)}\phi &:= \dot{\neg}\dot{(}\dot{\exists}\dot{x}\dot{)}\dot{\neg}\phi;\\
\dot{(}\phi\dot{\vee}\psi\dot{)} &:= \dot{\neg}\dot{(}\dot{\neg}\phi\dot{\wedge}\dot{\neg}\psi\dot{)};\\
\dot{(}\phi\dot{\rightarrow}\psi\dot{)} &:= \dot{\neg}\dot{(}\phi\dot{\wedge}\dot{\neg}\psi\dot{)};\\
\dot{(}\phi\dot{\leftrightarrow}\psi\dot{)} &:= \dot{(}\dot{(}\phi\dot{\rightarrow}\psi\dot{)}\dot{\wedge}\dot{(}\psi\dot{\rightarrow}\phi\dot{)}\dot{)}.
\end{aligned}$$

$\dot{0}\dot{x}$ is meant to be interpreted as $x = 0$ and $\dot{1}\dot{x}$ is meant to be interpreted as $x = 1$. Hence $\langle \mathbb{R}, +, \times, <, 0, 1\rangle$ is an $\mathcal{L}_{OF}$-structure and so is $\langle {}^*\mathbb{R}, {}^*+, {}^*\times, {}^*<, {}^*0, {}^*1\rangle$.

Definition 2.58. Define $\mathcal{R} := \langle \mathbb{R}, +, \times, <, 0, 1\rangle$ and ${}^*\mathcal{R} := \langle {}^*\mathbb{R}, {}^*+, {}^*<, {}^*0, {}^*1\rangle$.

Remark 2.59. Both $\mathcal{R}$ and ${}^*\mathcal{R}$ are $\mathcal{L}_{OF}$-structures.

Theorem 2.60. (Basic Transfer Principle). *If $\phi(x_0, \ldots, x_{n-1})$ is an $\mathcal{L}_{OF}$-formula with n free variables, then for all $r_0, \ldots, r_{n-1} \in \mathbb{R}$,*

$$\mathcal{R} \models \phi[r_0, \ldots, r_{n-1}] \quad \Leftrightarrow \quad {}^*\mathcal{R} \models \phi[{}^*r_0, \ldots, {}^*r_{n-1}].$$

Proof. The proof of this theorem is an immediate corollary of Łoś's Theorem 2.44.
($\Rightarrow$): $\phi[r_0, \ldots, r_{n-1}]$ is true in $\mathcal{R}$, then

$$\{i \in \mathbb{N} : \mathcal{R} \models \phi[\mathbf{r}_0(i), \ldots, \mathbf{r}_{n-1}(i)]\} = \{i \in \mathbb{N} : \mathcal{R} \models \phi[r_0, \ldots, r_{n-1}]\} = \mathbb{N},$$

where $\mathbf{r}_j$ is constant sequence $(r_j, r_j, \ldots)$ for all $j = 0, \ldots, n-1$. But $\mathbb{N} \in F$, therefore $\{i \in \mathbb{N} : \mathcal{R} \models \phi[\mathbf{r}_0(i), \ldots, \mathbf{r}_{n-1}(i)]\} \in F$ and by Łoś's Theorem, $\phi[{}^*r_0, \ldots, {}^*r_{n-1}]$ is true in $\mathcal{R}^{\mathbb{N}}/F = {}^*\mathcal{R}$.

($\Leftarrow$): Since $\phi[{}^*r_0, \ldots, {}^*r_{n-1}]$ is true in ${}^*\mathcal{R}$, by Łoś's Theorem we have:

$$\{i \in \mathbb{N} : \mathcal{R} \models \phi[\mathbf{r}_0(i), \ldots, \mathbf{r}_{n-1}(i)]\} = \{i \in \mathbb{N} : \mathcal{R} \models \phi[r_0, \ldots, r_{n-1}]\} \in F.$$

But the set $\{i \in \mathbb{N} : \mathcal{R} \models \phi[r_0, \ldots, r_{n-1}]\}$ is equal to $\mathbb{N} \in F$ if $\phi[r_0, \ldots, r_{n-1}]$ is true in $\mathcal{R}$, and is equal to $\emptyset \notin F$ if $\phi[r_0, \ldots, r_{n-1}]$ is not true in $\mathcal{R}$. Hence $\phi[r_0, \ldots, r_{n-1}]$ must be true in $\mathcal{R}$. □

For showing the usefulness and the power of the Transfer Principle, we give the proof of the Proposition 2.48 by using this machinery.

Proof of Proposition 2.48. We need to verify that the structure $\langle {}^*\mathbb{R}, {}^*+, {}^*\times, {}^*<, {}^*0, {}^*1 \rangle$ has the properties of an ordered field (for the properties of being an ordered field see page 63 in Bloch [12]). Note that, the property of being an ordered field is a sentence (i.e. a formula without free variables) in $\mathcal{L}_{OF}$. By the Transfer Principle this means that ${}^*\mathcal{R}$ is an ordered field if and only if $\mathcal{R}$ is. In detail:

1. (Commutative law for addition).

$$\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)} \quad \dot{(}\dot{x}\dot{+}\dot{y}\dot{=}\dot{y}\dot{+}\dot{x}\dot{)} \qquad \Leftrightarrow \qquad {}^*\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)} \quad \dot{(}\dot{x}\dot{+}\dot{y}\dot{=}\dot{y}\dot{+}\dot{x}\dot{)}.$$

2. (Associative law for addition).

$$\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)}\dot{(}\dot{\forall}\dot{z}\dot{)} \quad \dot{(}\dot{(}\dot{x}\dot{+}\dot{y}\dot{)}\dot{+}\dot{z}\dot{=}\dot{x}\dot{+}\dot{(}\dot{y}\dot{+}\dot{z}\dot{)}\dot{)}$$
$$\Updownarrow$$
$${}^*\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)}\dot{(}\dot{\forall}\dot{z}\dot{)} \quad \dot{(}\dot{(}\dot{x}\dot{+}\dot{y}\dot{)}\dot{+}\dot{z}\dot{=}\dot{x}\dot{+}\dot{(}\dot{y}\dot{+}\dot{z}\dot{)}\dot{)}.$$

3. (Identity law for addition).

$$\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)} \quad \dot{(}\dot{0}\dot{y} \dot{\rightarrow} \dot{x}\dot{+}\dot{y} \doteq \dot{x}\dot{)} \qquad \Leftrightarrow \qquad {}^*\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)} \quad \dot{(}\dot{0}\dot{y} \dot{\rightarrow} \dot{x}\dot{+}\dot{y} \doteq \dot{x}\dot{)}.$$

4. (Existence of an additive inverse).

$$\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\exists}\dot{y}\dot{)} \quad \dot{(}\dot{0}\dot{(}\dot{x}\dot{+}\dot{y}\dot{)}\dot{)} \qquad \Leftrightarrow \qquad {}^*\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\exists}\dot{y}\dot{)} \quad \dot{(}\dot{0}\dot{(}\dot{x}\dot{+}\dot{y}\dot{)}\dot{)}.$$

5. (Commutative law for multiplication).

$$\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)} \quad \dot{(}\dot{x}\dot{\times}\dot{y}\dot{=}\dot{y}\dot{\times}\dot{x}\dot{)} \qquad \Leftrightarrow \qquad {}^*\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)} \quad \dot{(}\dot{x}\dot{\times}\dot{y}\dot{=}\dot{y}\dot{\times}\dot{x}\dot{)}.$$

6. (Associative law for multiplication).

$$\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)}\dot{(}\dot\forall\dot z\dot{)} \quad \dot{(}\dot{(}\dot x\dot\times\dot y\dot{)}\dot\times\dot z\dot{=}\dot x\dot\times\dot{(}\dot y\dot\times\dot z\dot{)}\dot{)}$$
$$\Updownarrow$$
$$^*\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)}\dot{(}\dot\forall\dot z\dot{)} \quad \dot{(}\dot{(}\dot x\dot\times\dot y\dot{)}\dot\times\dot z\dot{=}\dot x\dot\times\dot{(}\dot y\dot\times\dot z\dot{)}\dot{)}.$$

7. (Identity law for multiplication).

$$\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)} \quad \dot{(}\dot 1\dot y \dot\to \dot x\dot\times\dot y \doteq \dot x\dot{)} \qquad \Leftrightarrow \qquad {}^*\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)} \quad \dot{(}\dot 1\dot y \dot\to \dot x\dot\times\dot y \doteq \dot x\dot{)}.$$

8. (Existence of a multiplicative inverse).

$$\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)} \quad \dot{(}\dot\neg\dot{(}\dot 0\dot x\dot{)} \dot\to \dot{(}\dot\exists\dot y\dot{)}\ \dot 1\dot{(}\dot x\dot\times\dot y\dot{)}\dot{)}$$
$$\Updownarrow$$
$$^*\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)} \quad \dot{(}\dot\neg\dot{(}\dot 0\dot x\dot{)} \dot\to \dot{(}\dot\exists\dot y\dot{)}\ \dot 1\dot{(}\dot x\dot\times\dot y\dot{)}\dot{)}.$$

9. (Distributive law).

$$\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)}\dot{(}\dot\forall\dot z\dot{)} \quad \dot{(}\dot x\dot\times\dot{(}\dot y\dot+\dot z\dot{)}\dot{=}\dot{(}\dot x\dot\times\dot y\dot{)}\dot+\dot{(}\dot x\dot\times\dot z\dot{)}\dot{)}$$
$$\Updownarrow$$
$$^*\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)}\dot{(}\dot\forall\dot z\dot{)} \quad \dot{(}\dot x\dot\times\dot{(}\dot y\dot+\dot z\dot{)}\dot{=}\dot{(}\dot x\dot\times\dot y\dot{)}\dot+\dot{(}\dot x\dot\times\dot z\dot{)}\dot{)}.$$

10. (Trichotomy law).

$$\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)} \qquad \dot{(}\dot x\dot<\dot y \ \dot\vee\ \dot x \doteq \dot y \ \dot\vee\ \dot x\dot>\dot y\dot{)}$$
$$\Updownarrow$$
$$^*\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)} \qquad \dot{(}\dot x\dot<\dot y \ \dot\vee\ \dot x \doteq \dot y \ \dot\vee\ \dot x\dot>\dot y\dot{)}.$$

11. (Transitive law).

$$\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)}\dot{(}\dot\forall\dot z\dot{)} \qquad \dot{(}\dot x\dot<\dot y \ \dot\wedge\ \dot y\dot<\dot z \ \dot\to\ \dot x\dot<\dot z\dot{)}$$
$$\Updownarrow$$
$$^*\mathcal{R} \models \dot{(}\dot\forall\dot x\dot{)}\dot{(}\dot\forall\dot y\dot{)}\dot{(}\dot\forall\dot z\dot{)} \qquad \dot{(}\dot x\dot<\dot y \ \dot\wedge\ \dot y\dot<\dot z \ \dot\to\ \dot x\dot<\dot z\dot{)}.$$

12. (Addition law for order).

$$\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)}\dot{(}\dot{\forall}\dot{z}\dot{)} \qquad \dot{(}\dot{x}\dot{<}\dot{y} \dot{\rightarrow} \dot{x}\dot{+}\dot{z}\dot{<}\dot{y}\dot{+}\dot{z}\dot{)}$$
$$\Updownarrow$$
$$^*\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)}\dot{(}\dot{\forall}\dot{z}\dot{)} \qquad \dot{(}\dot{x}\dot{<}\dot{y} \dot{\rightarrow} \dot{x}\dot{+}\dot{z}\dot{<}\dot{y}\dot{+}\dot{z}\dot{)}.$$

13. (Multiplication law for order).

$$\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)}\dot{(}\dot{\forall}\dot{z}\dot{)} \qquad \dot{(}\dot{x}\dot{<}\dot{y} \dot{\wedge} \dot{z}\dot{>}\dot{0} \dot{\rightarrow} \dot{x}\dot{\times}\dot{z}\dot{<}\dot{y}\dot{\times}\dot{z}\dot{)}$$
$$\Updownarrow$$
$$^*\mathcal{R} \models \dot{(}\dot{\forall}\dot{x}\dot{)}\dot{(}\dot{\forall}\dot{y}\dot{)}\dot{(}\dot{\forall}\dot{z}\dot{)} \qquad \dot{(}\dot{x}\dot{<}\dot{y} \dot{\wedge} \dot{z}\dot{>}\dot{0} \dot{\rightarrow} \dot{x}\dot{\times}\dot{z}\dot{<}\dot{y}\dot{\times}\dot{z}\dot{)}.$$

□

2.5 Superstructures and bounded ultrapowers

We will construct superstructures which will be used for introducing the nonstandard universe. We construct some superstructure over an arbitrary set S, by assuming that the relation $\in$ on this superstucture is the same as the ordinary element relation, except the elements of S are interpreted as *atoms* by $\in$, i.e. $v \notin s$ for all $s \in S$ and v is in the constructed superstructure over S (see Herzberg [30] and Anderson [4]).

Definition 2.61. Given a set S (with atom elements), *the superstructure* $V(S)$ *over* S is defined by

$$\begin{aligned} V_0(S) &= S, \\ V_n(S) &= V_{n-1}(S) \cup \mathcal{P}(V_{n-1}(S)), && \text{for all } n \in \mathbb{N} \\ V(S) &= \bigcup_{k\in\mathbb{N}} V_k(S). \end{aligned}$$

Example 2.4. Let $V(S)$ be a superstructure such that S contains $\mathbb{N}$. The number 7 and the set $\{7\}$ are in $V_1(S)$. The number 7, the set $\{7\}$, and the set of all finite subsets of $\mathbb{N}$ are in $V_2(S)$.

In general, we have the following remark for the superstructure $V(S)$ over an arbitrary set S:

Remark 2.62. (1) $V_0(S) \subseteq V_1(S) \subseteq \cdots \subseteq V_{n-1}(S) \subseteq \cdots$.

(2) $V(S)$ consists of atoms and sets. That is,

$$\forall s \in V_0(S) \quad [s \neq \emptyset \quad \wedge \quad (\forall v \in V(S),\, v \notin s)]\,,$$

and all other members of $V(S)$ are sets.

(3) $V_n(S) = S \cup \mathcal{P}(V_{n-1}(S))$.

(4) $V_{n-1}(S) \in V_n(S)$. Hence $V_{n-1}(S) \in V(S)$, and in particular $S \in V(S)$.

(5) $a \in A \in V_n(S)$ implies $a \in V_{n-1}(S)$.

(6) If $a, b \in V_{n-1}(S)$, then $\{a, b\} \in V_n(S)$.

(7) If $A, B \in V_{n-1}(S)$, then $A \cup B \in V_n(S)$.

(8) $A \in V_{n-1}(S)$ implies $\mathcal{P}(A) \in V_{n+1}(S)$.

(9) Each $V_n(S)$ is transitive, i.e., each element of $V_n(S)$ which is not an atom, is a subset of $V_n(S)$.

Properties (6)–(8) ensure that $V(S)$ is a universe, and by (4) it is a universe over S. In fact, $V(S)$ is the smallest universe containing S, that is, if any universe V' has the property $S \in V'$, then $V(S) \subseteq V'$.

Remark 2.63. The superstructure $V(S)$ consists of all mathematical objects. The atoms are in $V_0(S)$; the ordered pairs $(x, y) \in S \times S$ belongs to $V_2(S)$, since they can be perceived as sets of the type $\{\{x\}, \{x, y\}\}$; the functions $f : S \to S$, and more generally, the relations in S are subsets of $V_2(S)$ and hence, belong to $V_3(S)$. If $\mathscr{O}$ is a topology on S, then $\mathscr{O} \subseteq \mathcal{P}(S)$ and hence $\mathscr{O}$ belongs to $V_2(S)$. Where $S = X \cup \mathbb{R}$, the algebraic operations in S are perceived as subsets of $S \times S \times S$ and hence also belong to $V(S)$.

Definition 2.64. A sequence $A = (A(0), A(1), \ldots)$ of elements of $V(S)$ is *bounded* if there exists a fixed $n \in \mathbb{N}$ such that for every $i \in \mathbb{N}$, $A(i) \in V_n(S)$.

Definition 2.65. Let F be a non-principal ultrafilter on $\mathbb{N}$. Two bounded sequences A and B are equivalent with respect to F if and only if $\{i \in \mathbb{N} : A(i) = B(i)\} \in F$, and we denote it by $A \sim_F B$. We let $[A]_F$ be the equivalence class of A and define the *bounded ultrapower* by

$$V(S)^{\mathbb{N}}/F := \{[A]_F : A \text{ is a bounded sequence in } V(S)\}.$$

We define the membership relation $\in_F$ in the bounded ultrapower as follows:

Definition 2.66. Let F be a non-principal ultrafilter on $\mathbb{N}$ and suppose $[A]_F$ and $[B]_F$ are equivalence classes of the sequences A and B in $V(S)$. Then

$$[A]_F \in_F [B]_F \quad \Leftrightarrow \quad \{i \in \mathbb{N} : A(i) \in B(i)\} \in F.$$

As in Section 2.4, we need a formal language for $V(S)$ to state *Łoś's Theorem* and the *Transfer Principle*. The language $\mathcal{L}_{V(S)}$ consists of the symbols

$$\{\dot{x}, \dot{y}, \dot{z}, \dot{x}_0, \dot{x}_1, \ldots, \dot{y}_0, \dot{y}_1, \ldots\} \cup \{\dot{\in}\} \cup \{\dot{=}, \dot{\exists}, \dot{\wedge}, \dot{\neg}, \dot{(}, \dot{)}\},$$

where $\dot{x}, \dot{y}, \dot{z}, \dot{x}_0, \dot{x}_1, \ldots, \dot{y}_0, \dot{y}_1, \ldots$ are *variables*, $\dot{\in}$ is the *set membership*, $\dot{=}$ is the *equality symbol*, $\dot{\exists}$ is the *existential quantifier*, $\dot{\wedge}$ and $\dot{\neg}$ are *logical connectives*, and $\dot{(}$ and $\dot{)}$ are *parentheses*.

A string is an *atomic formula* if and only if it is of the form $\dot{x} \doteq \dot{y}$ or $\dot{x} \,\dot{\in}\, \dot{y}$. The set of *formulae* is the smallest subset of strings for which every atomic formula is a formula of $\mathcal{L}_{V(S)}$, and if ϕ, ψ are formulae of $\mathcal{L}_{V(S)}$ and $\dot{x}$ is a variable, then

$$\dot{\neg}\phi, \quad \dot{(}\phi \dot{\wedge} \psi\dot{)}, \quad \dot{(}\dot{\exists}\dot{x}\dot{)}\phi,$$

are formulae of $\mathcal{L}_{V(S)}$. We can achieve the other formulae with abbreviations (see Section 2.2). The atomic formulae $\dot{x} \doteq \dot{y}$ and $\dot{x} \,\dot{\in}\, \dot{y}$ are interpreted as $x = y$, $x \in y$ for every variables $x, y \in S$, respectively. Since the logical symbols have a fixed meaning over any domain, this means that every formula ϕ of $\mathcal{L}_{V(S)}$ has an interpretation in $V(S)$.

Łoś's Theorem extends to the bounded ultrapower $V(S)^{\mathbb{N}}/F$ by a very similar proof.

Theorem 2.67. (Łoś's Theorem for Bounded Ultrapowers). *If F is a (non-principal) ultrafilter on $\mathbb{N}$ and ϕ is an $\mathcal{L}_{V(S)}$-formula with bounded quantifiers and n free variables, then for all bounded sequences $A_0, \ldots, A_{n-1} \in V(S)^{\mathbb{N}}$,*

$$V(S)^{\mathbb{N}}/F \models \phi[[A_0]_F, \ldots, [A_{n-1}]_F] \Leftrightarrow \{i \in \mathbb{N} : V(S) \models \phi[A_0(i), \ldots, A_{n-1}(i)]\} \in F.$$

Proof. The idea of the proof is that we modify Łoś's Theorem 2.44 for bounded, rather than ordinary, ultrapowers. We again proceed by induction on the length of ϕ. At first, suppose ϕ is an atomic formula of the form $\phi = \dot{x} \doteq \dot{y}$. Then for some $k, m \in \mathbb{N}$, we have

$$\begin{aligned} V(S)^{\mathbb{N}}/F \models [A_k]_F \doteq [A_m]_F &\Leftrightarrow [A_k]_F = [A_m]_F \\ &\Leftrightarrow \{i \in \mathbb{N} : A_k(i) = A_m(i)\} \in F. \quad \text{(by definition of } \sim_F) \end{aligned}$$

Since A_k, A_m are bounded sequences, there exists a fixed $n \in \mathbb{N}$ such that for every $i \in \mathbb{N}$, $A_k(i) \in V_n(S) \subseteq V(S)$ and $A_m(i) \in V_n(S) \subseteq V(S)$. Hence

$$\{i \in \mathbb{N} : A_k(i) = A_m(i)\} \in F \Leftrightarrow \{i \in \mathbb{N} : V(S) \models A_k(i) \doteq A_m(i)\} \in F.$$

Therefore, the theorem is true for the atomic formula of the form $\phi = \dot{x} \doteq \dot{y}$. Now suppose that ϕ is an atomic formula of the form $\phi = \dot{x} \,\dot{\in}\, \dot{y}$. Then for some $k, m \in \mathbb{N}$, we have

$$V(S)^{\mathbb{N}}/F \models [A_k]_F \,\dot{\in}\, [A_m]_F \;\Leftrightarrow\; [A_k]_F \in_F [A_m]_F$$
$$\Leftrightarrow \{i \in \mathbb{N} : A_k(i) \in A_m(i)\} \in F. \quad \text{(by definition of } \in_F\text{)}$$

Again, since A_k and A_m are bounded sequences,

$$\{i \in \mathbb{N} : A_k(i) \in_F A_m(i)\} \in F \;\Leftrightarrow\; \{i \in \mathbb{N} : V(S) \models A_k(i) \,\dot{\in}\, A_m(i)\} \in F.$$

Hence, the theorem is true for the atomic formula of the form $\phi = \dot{x} \,\dot{\in}\, \dot{y}$.

For induction steps, the proof of the cases $\phi = \dot{\neg}\psi$ and $\phi = \psi \dot{\wedge} \chi$ for some $\mathcal{L}_{V(S)}$-formula ψ and χ, is exactly the same with the proof of Łoś's Theorem 2.44. It remains to consider the case where ϕ is of the form $\dot{(}\dot{\exists}\dot{x} \,\dot{\in}\, \dot{y}\dot{)}\, \psi(\dot{x}_0, \ldots, \dot{x}_{n-1}, \dot{x})$. We want to show that

$$V(S)^{\mathbb{N}}/F \models \dot{(}\dot{\exists}\dot{x} \,\dot{\in}\, [A_m]_F\dot{)}\, \psi\,[[A_0]_F, \ldots, [A_{n-1}]_F, \dot{x}]$$
$$\Leftrightarrow \{i \in \mathbb{N} : V(S) \models \dot{(}\dot{\exists}\dot{x} \,\dot{\in}\, A_m(i)\dot{)}\, \psi[A_0(i), \ldots, A_{n-1}(i), \dot{x}]\} \in F.$$

Suppose $s = \{i \in \mathbb{N} : V(S) \models \dot{(}\dot{\exists}\dot{x} \,\dot{\in}\, A_m(i)\dot{)}\, \psi[A_0(i), \ldots, A_{n-1}(i), \dot{x}]\} \in F$. Then for each $i \in s$, there exists in $V(S)$ an element $C(i) \in A_m(i)$ such that

$$V(S) \models \psi[A_0(i), \ldots, A_{n-1}(i), C(i)].$$

Let $h : \mathbb{N} \to S$ be such that $h(i) \in A_m(i)$ for all $i \in \mathbb{N}$. Define the function $g : \mathbb{N} \to V(S)$ by

$$g(i) := \begin{cases} C(i), & \text{for } i \in s \\ h(i), & \text{for } i \notin s \end{cases}.$$

Since A_m is a bounded sequence, there exists some $n \in \mathbb{N}$ such that

$$\{i \in \mathbb{N} : A_m(i) \in V_n(S)\} \in F,$$

and by transitivity of $V_n(S)$,

$$\{i \in \mathbb{N} : g(i) \in V_n(S)\} \supseteq \{i \in \mathbb{N} : A_m(i) \in V_n(S)\} \in F.$$

But F is closed under supersets, hence $\{i \in \mathbb{N} : g(i) \in V_n(S)\} \in F$ and therefore $[g]_F \in V(S)^{\mathbb{N}}/F$. Moreover

$$\{i \in \mathbb{N} : V(S) \models \psi[A_0(i), \ldots, A_{n-1}(i), g(i)]\} \supseteq s \in F.$$

Thus $\{i \in \mathbb{N} : V(S) \models \psi[A_0(i), \ldots, A_{n-1}(i), g(i)]\} \in F$. By induction hypothesis,

$$V(S)^{\mathbb{N}}/F \models \psi[[A_0]_F, \ldots, [A_{n-1}]_F, [g]_F].$$

Therefore

$$V(S)^{\mathbb{N}}/F \models \dot{(}\dot{\exists}\dot{x} \mathrel{\dot{\in}} [A_m]_F\dot{)}\, \psi[[A_0]_F, \ldots, [A_{n-1}]_F, \dot{x}].$$

Now suppose that $V(S)^{\mathbb{N}}/F \models \dot{(}\dot{\exists}\dot{x} \mathrel{\dot{\in}} [A_m]_F\dot{)}\, \psi\, [[A_0]_F, \ldots, [A_{n-1}]_F, \dot{x}]$. This means that there exists $[g]_F \in_F [A_m]_F$ such that

$$V(S)^{\mathbb{N}}/F \models \psi\, [[A_0]_F, \ldots, [A_{n-1}]_F, [g]_F]\,.$$

By induction hypothesis,

$$\{i \in \mathbb{N} : V(S) \models \psi[A_0(i), \ldots, A_{n-1}(i), g(i)]\} \in F.$$

But

$$\begin{aligned}\{i \in \mathbb{N} : V(S) \models \dot{(}\dot{\exists}\dot{x} \mathrel{\dot{\in}} A_m(i)\dot{)}\, \psi[A_0(i), \ldots, A_{n-1}(i), \dot{x}]\}& \\ \supseteq \{i \in \mathbb{N} : V(S) \models \psi[A_0(i), \ldots, A_{n-1}(i), g(i)]\} \in F.&\end{aligned}$$

Since F is closed under supersets,

$$\{i \in \mathbb{N} : V(S) \models \dot{(}\dot{\exists}\dot{x} \mathrel{\dot{\in}} A_m(i)\dot{)}\, \psi[A_0(i), \ldots, A_{n-1}(i), \dot{x}]\} \in F,$$

and this completes the proof of the theorem. □

2.6 Nonstandard universe

In Section 2.4, we have shown that $^*\mathbb{R}$ is a proper extension of $\mathbb{R}$ which already contains infinitesimals and infinitely large numbers. But it is an insufficient setting for many applications of nonstandard analysis. We need an *extended universe* that in addition to numbers and functions, also contains sets of functions, sets of spaces of functions. Therefore we require an enlargement of the whole superstructure over the reals, that is $V(\mathbb{R})$. However, for the sake of generality, we consider the case of the superstructure $V(S)$ over an arbitrary set S.

As we have introduced in Section 2.5, the superstructure over S is obtained by iterating the power-set operator countably many times. The nonstandard universe will be constructed by assuming an extension $^*S$ of S and assuming an embedding

$$^* : V(S) \to V(^*S), \tag{2.9}$$

with the properties similar to the embedding (2.3). At first, we have the following principle:

Extension Principle. ${}^*s = [\mathbf{s}]_F$ for every $s \in S$, where $[\mathbf{s}]_F \in S^{\mathbb{N}}/F$ is the equivalence class of the constant sequence $\mathbf{s} = (s, s, \ldots) \in S^{\mathbb{N}}$ with respect to the non-principal ultrafilter F.

We shall demonstrate that the superstructure embedding (2.9) satisfies the Transfer Principle. We do this in two stages. First, we map the superstructure $V(S)$ into the bounded ultrapower $V(S)^{\mathbb{N}}/F$, which has been introduced in Definition 2.65. Then we map this bounded ultrapower into the superstructure $V({}^*S)$ in such a way that the embedding 2.9 satisfies the Transfer Principle.

- *Embedding $V(S)$ into $V(S)^{\mathbb{N}}/F$:* We consider the following embedding for every $A \in V(S)$:
$$i : V(S) \to V(S)^{\mathbb{N}}/F,$$
such that $i(A) = [\mathbf{A}]_F$, where $\mathbf{A} = (A, A, \ldots)$ is a bounded constant sequence in $V(S)$ and $[\mathbf{A}]_F$ is the equivalence class of $\mathbf{A}$ with respect to the non-principal ultrafilter F.

- *Embedding $V(S)^{\mathbb{N}}/F$ into $V({}^*S)$:* We construct the following embedding:
$$j : V(S)^{\mathbb{N}}/F \to V({}^*S),$$
such that (i) j is identity on ${}^*S$ and (ii) if $[\mathbf{A}]_F \notin {}^*S$, then
$$j([\mathbf{A}]_F) = \{j([\mathbf{B}]_F) \,:\, [\mathbf{B}]_F \in_F [\mathbf{A}]_F\},$$
where $\in_F$ is defined in Definition 2.66.

Remark 2.68. The last embedding mapped the relation $\in_F$ in the ultrapower into the ordinary membership relation in $V({}^*S)$.

Remark 2.69. We can define j by induction. Let
$$V_k(S)^{\mathbb{N}}/F = \{[A]_F \,:\, A \text{ is a sequence in } V_k(S)\}.$$
Then the bounded ultrapower is the union of the chain
$${}^*S = V_0(\mathbb{R})^{\mathbb{N}}/F \subseteq \ldots \subseteq V_k(S)^{\mathbb{N}}/F \subseteq \ldots .$$
Then for $k = 0$, j must be the identity. If $[A]_F \in V_k(S)^{\mathbb{N}}/F$ and $[A]_F \notin {}^*S$, we set $j([A]_F) = \{j([B]_F) \,:\, [B]_F \in_F [A]_F\}$. This makes sense, since if $[B]_F \in_F [A]_F$, it follows from Definition 2.66 that $\{i \in \mathbb{N} \,:\, B_i \in V_{k-1}(S)\} \in F$, i.e., $[B]_F \in V_{k-1}(S)^{\mathbb{N}}/F$, which means that $j([B]_F)$ is defined at the previous stage of the inductive construction.

Definition 2.70. The composition of i and j, denoted by $^* : j \circ i$, is an embedding of the structure $V(S)$ into $V(^*S)$, where for every $A \in V(S)$, $^*A = j(i(A))$. We call $V(^*S)$ the *extended nonstandard universe.*

Now, we give the Transfer Principle:

Theorem 2.71. (Transfer Principle). *If $\phi(x_0, \ldots, x_{n-1})$ is an $\in$-formula with bounded quantifiers and n free variables, then for all $A_0, \ldots, A_{n-1} \in V(S)$,*

$$V(S) \models \phi[A_0, \ldots, A_{n-1}] \quad \Leftrightarrow \quad V(^*S) \models \phi[^*A_0, \ldots, {}^*A_{n-1}].$$

Proof. According to Łoś's Theorem for Bounded Ultrapowers (Theorem 2.67), for any $[A_0]_F, \ldots, [A_{n-1}]_F \in V(S)^{\mathbb{N}}/F$, we have

$$V(S)^{\mathbb{N}}/F \models \phi\,[[A_0]_F, \ldots, [A_{n-1}]_F] \quad \Leftrightarrow \quad \{i \in \mathbb{N} : V(S) \models \phi\,[A_0(i), \ldots A_{n-1}(i)]\} \in F, \tag{2.10}$$

from which transfer follows between $V(S)$ and $V(S)^{\mathbb{N}}/F$, exactly as in 2.60.

But we are looking for the transfer between $V(S)$ and $V(^*S)$. In order to prove this we need to replace 2.10 by

$$V(S)^{\mathbb{N}}/F \models \phi[j([A_0]_F), \ldots, j([A_{n-1}]_F)] \Leftrightarrow \{i \in \mathbb{N} : V(S) \models \phi[[A_0]_i, \ldots [A_{n-1}]_i]\} \in F. \tag{2.11}$$

This is also an immediate extension which follows from the fact that every element $j([A]_F)$ in $V(^*S)$ is of the form $j([B]_F)$ for some $[B]_F$ in $V(S)^{\mathbb{N}}/F$ (see the construction of the j-map). Therefore the Extended Transfer Principle 2.71 follows by the same proof as in 2.60. □

Remark 2.72. If $S = \mathbb{R}$, all properties of $^*\mathbb{R}$ discussed in Section 2.4 hold for $^*\mathbb{R}$ in $V(^*\mathbb{R})$. Elements of $^*\mathbb{R}$ are either *finite* or *infinite*, and every finite element in $^*\mathbb{R}$ has a unique *standard part* in $\mathbb{R}$.

Definition 2.73. Let $A \in V(^*S)$, then

- A is called *standard* if $A = {}^*B$ for some $B \in V(S)$,
- A is called *internal* if $A \in {}^*B$ for some $B \in V(S)$.
- A is called *external* if A is not internal.

Remark 2.74. $A \in V(^*S)$ is internal if and only if there exists $m \in \mathbb{N}$ such that $A \in {}^*V_m(S)$.

Proof. Suppose there exists $m \in \mathbb{N}$ such that $A \in {}^*V_m(S)$. Since $V_m(S) \in V(S)$, A is internal. Conversely, suppose that A is internal. By the previous definition, there exists $m \in \mathbb{N}$ such that $A \in {}^*B$ and $B \in V_{m+1}(S)$. Hence $A \in {}^*V_m(S)$. This is true, because $\forall m \in \mathbb{N}$ and $\forall x \in y \in {}^*V_{m+1}(S)$, $x \in {}^*V_m(S)$. □

Remark 2.75. $\mathbb{N}$ is an external set.

Proof. Suppose that $\mathbb{N}$ is internal, which means that $\mathbb{N} \in {}^*\mathcal{P}(\mathbb{N})$. Let

$$\forall A \in \mathcal{P}(\mathbb{N}) \quad (\mathbb{N} \backslash A \in \mathcal{P}(\mathbb{N})).$$

By the Transfer Principle,

$$\forall A \in {}^*\mathcal{P}(\mathbb{N}) \quad ({}^*\mathbb{N} \backslash A \in {}^*\mathcal{P}(\mathbb{N})),$$

and hence the set ${}^*\mathbb{N} \backslash \mathbb{N}$ is also internal. Again, by using the Transfer Principle for

$$\forall A \in \mathcal{P}(\mathbb{N}) \quad (A \neq \emptyset \Rightarrow (\exists m \in A)(\forall k \in A)(m \leq k)),$$

there is a first element m in ${}^*\mathbb{N} \backslash \mathbb{N}$. Therefore, $m - 1$ is the last element of $\mathbb{N}$, which is impossible. Hence $\mathbb{N}$ is external. □

Proposition 2.76. (Internal Definition Principle). *Let $B, A_0, A_1, \ldots, A_{n-1}$ be internal sets in $V({}^*S)$ and let $\phi(x, x_0, x_1, \ldots, x_{n-1})$ be an $\in$-formula with bounded quantifiers and $n + 1$ free variables, then the set*

$$\{y \in B \,:\, V({}^*S) \models \phi\,[y, A_0, A_1, \ldots, A_{n-1}]\},$$

is internal.

Proof. Since $B, A_0, A_1, \ldots, A_{n-1}$ are internal sets in $V({}^*S)$, there exists some integer m such that $B, A_0, A_1, \ldots, A_{n-1} \in {}^*V_m(S)$. In $V(S)$ we have the following true sentence:

$$\forall x, x_0, x_1, \ldots, x_{n-1} \in V_m(S) \quad \exists z \in V_{m+1}(S) \quad (z = \{t \in x \,:\, \phi\,(t, x_0, x_1, \ldots, x_{n-1})\}),$$

where $z = \{t \in x \,:\, \phi\,(t, x_0, x_1, \ldots, x_{n-1})\}$ is an abbreviation of the formula

$$\forall t (t \in z \Leftrightarrow t \in x \wedge \phi(t, x_0, x_1, \ldots, x_{n-1})),$$

which implies

$$\forall t \in V_{m+1}(S)\ (t \in z \Leftrightarrow t \in x \wedge \phi(t, x_0, x_1, \ldots, x_{n-1}))\,.$$

By using the Transfer Principle we have

$$\begin{aligned} &\forall x, x_0, x_1, \ldots, x_{n-1} \in {}^*V_m(S) \quad \exists z \in {}^*V_{m+1}(S) \\ &\forall t \in {}^*V_{m+1}(S) \quad (t \in z \Leftrightarrow t \in x \wedge \phi(t, x_0, x_1, \ldots, x_{n-1}))\,, \end{aligned}$$

is true in $V(^*S)$. Since $B, A_1, \ldots, A_n \in {}^*V_m(S)$,

$$\forall y \in {}^*V_{m+1}(S) \quad (y \in z \Leftrightarrow y \in B \wedge V(^*S) \models \phi[y, A_0, A_1, \ldots, A_{n-1})] .$$

But $z \in {}^*V_{m+1}(S)$ and by the transitivity of ${}^*V(S)$, $z \subseteq {}^*V_{m+1}(S)$. Hence $z \cap {}^*V_{m+1}(S) = z$. Which means that

$$\{y \in B \,:\, V\,(^*S) \models \phi\,[y, A_0, A_1, \ldots, A_{n-1}]\},$$

is in ${}^*V_{m+1}(S)$, and therefore it is internal. □

Example 2.5. Since ${}^*\mathbb{N}$ is internal, any $n \in {}^*\mathbb{N}$ is internal. Then the set

$$\{v \in {}^*\mathbb{N} \,:\, v > n\},$$

is also an internal set.

Proposition 2.77. (Overspill Principle). *Let $A \subseteq {}^*\mathbb{R}$ be a non-empty internal set. If A contains arbitrary large finite positive numbers, then it also contains an infinite number.*

Proof. (First proof). If A is unbounded, then A contains at least one infinite element and so the assertion is obvious. Otherwise, let

$$X = \{m \in {}^*\mathbb{N} \,:\, m \text{ is an upper bound of } A\}.$$

By the Internal Definition Principle, X is a non-empty internal subset of ${}^*\mathbb{N}$, with a least element M (every non-empty subset of $\mathbb{N}$ has a least element, by the Transfer Principle, every non-empty internal subset of ${}^*\mathbb{N}$ has a least element). Since A contains arbitrary large finite positive numbers, M must be infinite. Indeed, if there is no $x \in A$ so that $M - \epsilon \leq x \leq M$ for some positive ϵ, then $M - \epsilon$ is the least upper bound of A, which is a contradiction. So there is some $x \in A$ such that $M - \epsilon \leq x \leq M$, and hence this completes the proof of the theorem. □

Proof. (Second proof). In the first proof we have shown that A is a non-empty internal subset of ${}^*\mathbb{N}$. Furthermore, by hypothesis, we know that $\exists N \in \mathbb{N}$ such that

$$\{N, N + 1, \ldots\} \subseteq A.$$

We want to show that there exists $M \in {}^*\mathbb{N}\backslash\mathbb{N}$ such that $M \in A$. For proving this by contraposition, suppose that $\nexists M \in {}^*\mathbb{N}\backslash\mathbb{N}$ such that $M \in A$. Therefore $A = \{N, N + 1, \ldots\}$ is internal, and hence $\{0, 1, \ldots, N-1\} \cup \{N, N+1, \ldots\} = \mathbb{N}$ is internal, which is a contradiction with the Remark 2.75 □

In the future, we will deal with some normed linear spaces, especially Banach spaces, over real numbers, and will prove some theorems by using the machinery that we have described in this chapter. Therefore, we need to construct the superstructure $V(X\cup\mathbb{R})$, where X is a Banach space over $\mathbb{R}$, by assuming that the elements of X and $\mathbb{R}$ are interpreted as atoms. Then we construct a bounded ultrapower of $V(X\cup\mathbb{R})$ by collecting the equivalence classes of sequences in $V(X\cup\mathbb{R})$, that are bounded in the superstructure hierarchy, using the non-principal ultrafilter F on $\mathbb{N}$. As we have seen in the construction of $^* : j\circ i$, the ultrapower construction can easily be adapted to construct an embedding

$$^* : V(X\cup\mathbb{R}) \to V(^*X \cup {}^*\mathbb{R}),$$

where $^*X$ is a Banach space over $^*\mathbb{R}$ satisfying:

- *Extension Principle*: $^*X$ and $^*\mathbb{R}$ are proper extensions of X and $\mathbb{R}$, respectively, and $^*x = [\mathbf{x}]_F$ for every $x \in X\cup\mathbb{R}$, and
- *Transfer Principle*: If $\phi(v_0, \cdots, v_{n-1})$ is an $\in$-formula with bounded quantifiers and n free variables, then for every $A_0, \cdots, A_{n-1} \in V(X\cup\mathbb{R})$,

$$V(X\cup\mathbb{R}) \models \phi[A_0, \cdots, A_{n-1}] \quad \Leftrightarrow \quad V(^*X\cup {}^*\mathbb{R}) \models \phi[{}^*A_0, \cdots, {}^*A_{n-1}].$$

Convention 2.78. If $x \in X\cup\mathbb{R}$, we write x for $^*x$.

2.7 Saturation and topology

We shall introduce a new principle of nonstandard analysis, *saturation*, which is very important when we are dealing with mathematical objects such as Banach spaces and topological spaces.

Proposition 2.79. (Countable Saturation Principle). *If $A_0 \supseteq A_1 \supseteq \dots$ is a decreasing countable sequence of non-empty internal sets in $V(^*S)$, then*

$$\bigcap_{i\in\mathbb{N}} A_i \neq \emptyset.$$

Proof. For any $i\in\mathbb{N}$, since A_i is internal, it is of the form $A_i = j(A_i')$, where

$$A_i' = [(A_0^i, A_1^i, \ldots)]_F.$$

We may assume that each $A_k^i \subseteq V_n(S)$ for some fixed $n\in\mathbb{N}$. For $k\geq 0$, let

$$I_k := \{i \geq k \,:\, A_i^0 \supseteq A_i^1 \supseteq \ldots \supseteq A_i^k \neq \emptyset\}.$$

Then $I_0 = \mathbb{N}$ and for all $k \in \mathbb{N}$, $I_k \in F$ and $I_{k-1} \supseteq I_k$. Also we have $\bigcap_{k\in\mathbb{N}} I_k = \emptyset$. This implies that

$$m(i) := \max\{m \,:\, i \in I_m\},$$

is well-defined for each $i \in \mathbb{N}$. Let B_i be some element in $A_i^{m(i)}$. We shall prove that $B = [(B_0, B_1, \ldots)]_F \in_F A'_k$ for every $k \geq 0$. Note that $i \in I_k$ implies that $m(i) \geq k$, and hence $B_i \in A_i^{m(i)} \subseteq A_i^k$. Thus

$$\{i \in \mathbb{N} \,:\, B_i \in A_i^k\} \supseteq I_k \in F,$$

and since F is an ultrafilter, $\{i \in \mathbb{N} \,:\, B_i \in A_i^k\} \in F$. Therefore $B = [(B_0, B_1, \ldots)]_F \in_F A'_k$ and according to the construction of j-map, $j(A'_k) = j(B) = A_k$. Since k is arbitrary, the proof is finished. □

Definition 2.80. A family $\mathcal{F}$ of subsets of some set S has the *finite intersection property (fip)* if for every finite set of elements $F_0, F_1, \ldots, F_{n-1}$ in $\mathcal{F}$, we have

$$F_0 \cap F_1 \cap \ldots \cap F_{n-1} \neq \emptyset.$$

Definition 2.81. (κ-Saturation Principle). Let κ be an infinite cardinal. * is κ-saturated if and only if for every collection of internal sets $(A_i)_{i\in I}$ in $V({}^*S)$ with $|I| < \kappa$, if $(A_i)_{i\in I}$ has the finite intersection property, then the intersection $\bigcap_{i\in I} A_i$ is non-empty, i.e. contains some internal object. Note that, if $\kappa = \aleph_1$, the first uncountable cardinal, then the κ-saturation principle is Countable Saturation Principle 2.79.

Remark 2.82. Chang and Keisler in [14] (Chapter 5) have proved the existence of the κ-saturated embeddings.

Now, we want to introduce the monad of an element of some topological space X and then give some results about the compactness of X. Later we will talk about the continuity of a function on X.

Definition 2.83. Let $(X, \mathscr{O})$ be a topological space and for $x \in X$, $\mathscr{O}_x$ be the family of open sets containing x. The *monad* of $x \in X$ is given by

$$\mu(x) := \bigcap\{{}^*\mathcal{O} \,:\, \mathcal{O} \in \mathscr{O}_x\}.$$

Remark 2.84. If $y \in {}^*X$ and $y \in \mu(x)$ for some $x \in X$, then we write $y \approx_X x$.

Definition 2.85. Suppose $(X, \mathscr{O})$ is a topological space. An element $x \in {}^*X$ is *near-standard* if and only if $x \in \mu(y)$ for some $y \in X$.

Remark 2.86. If $(X, \mathscr{O})$ is a Hausdorff space, $x \in {}^*X$ and $x \in \mu(y)$ for some $y \in X$, then y is unique and in this case we write $y = {}^\circ x$.

Proof. The proof of this remark is an immediate corollary of the fact that, $(X, \mathscr{O})$ is a Hausdorff space if and only if for every $x, y \in X$ with $x \neq y$, $\mu(x) \cap \mu(y) = \emptyset$. The reader can find the proof of this fact at page 48 in Albeverio et al. [2]. $\square$

Proposition 2.87. *Suppose $(X, \mathscr{O})$ is a topological space. If X is compact, then every element in ${}^*X$ is nearstandard.*

Proof. Suppose X is compact and that there is some $y \in {}^*X$ which is not nearstandard. Then for every $x \in X$, there exists $\mathscr{O}_x$ with $x \in \mathscr{O}_x \in \mathscr{O}$ and $y \notin {}^*\mathscr{O}_x$. Therefore $\{\mathscr{O}_x : x \in X\}$ is an open cover of X. Let $\{\mathscr{O}_{x_0}, \ldots, \mathscr{O}_{x_{n-1}}\}$ be a finite subcover of X for $n \in \mathbb{N}$. Hence

$$\bigcup_{i=0}^{n-1} {}^*\mathscr{O}_{x_i} = {}^*\left(\bigcup_{i=0}^{n-1} \mathscr{O}_{x_i}\right) = {}^*X,$$

so $y \notin {}^*X$, which is a contradiction, and this completes the proof of the proposition. $\square$

We have the following definition from Diener and Diener [18].

Definition 2.88. Suppose $(X, \mathscr{O})$ and $(Y, \mathscr{T})$ are topological spaces. $f : {}^*X \to {}^*Y$ is *S-continuous* in $x \in {}^*X$ if and only if

$$\forall y \in {}^*X \quad (x \approx_X y \Rightarrow f(x) \approx_Y f(y)).$$

f is *S-continuous on* ${}^*X$ if and only if for all nearstandard $x \in {}^*X$, f is S-continuous in x.

Proposition 2.89. *If $f : X \to \mathbb{R}$ is continuous, then ${}^*f : {}^*X \to {}^*\mathbb{R}$ is S-continuous.*

Proof. Choose a nearstandard $\bar{x} \in {}^*X$. Let $x \approx \bar{x}$ for some $x \in X$. Since f is continuous, it is also continuous in x, and we can write

$$(\forall \epsilon \in \mathbb{R}_+)\,(\exists \delta \in \mathbb{R}_+)\,\left((\forall y \in X)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon)\right).$$

Transfer Principle can be applied to $(\forall y \in X)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$, and for every $\epsilon \in \mathbb{R}_+$ there exists $\delta \in \mathbb{R}_+$ such that

$$(\forall y \in {}^*X)(|x - y| < \delta \Rightarrow |{}^*f(x) - {}^*f(y)| < \epsilon).$$

Now suppose that $y \in {}^*X$ and $y \approx \bar{x}$. Thus, since $x \approx \bar{x}$, we have $y \approx x$; hence $|x-y| < \delta$ for all $\delta \in \mathbb{R}_+$. Then $|{}^*f(x) - {}^*f(y)| < \epsilon$ for all $\epsilon \in \mathbb{R}_+$. Therefore ${}^*f(x) \approx {}^*f(y)$. In particular, this holds for $y = \bar{x}$. So ${}^*f(x) \approx {}^*f(y)$ and ${}^*f(x) \approx {}^*f(\bar{x})$. Hence ${}^*f(\bar{x}) \approx {}^*f(y)$ for all $y \approx \bar{x}$. $\square$

Chapter 3

Theorem on microeconomic foundations of representative agent models

This chapter builds on a recent proposal for microeconomic foundations of "representative agents". Herzberg in [31] constructed a representative utility function for finite-dimensional social decision problems and since the decision problems of macroeconomic theory are typically infinite-dimensional, Herzberg's original result is insufficient for many applications. We therefore generalise his result by allowing the social alternatives to belong to a general Banach space.

3.1 The model and formulation

We are concerned with a social decision problem, therefore we need a model for introducing population, alternatives and utility functions. We use the following model:

3.1.1 Individuals and social alternatives

Let N be a set of *individuals*. Subsets of N are called *coalitions*. We fix some subset $\mathcal{D}$ of the power-set of N and call the elements of $\mathcal{D}$ *potentially decisive coalitions*.

We also let $\mathcal{C}$ be a set of *social alternatives*. For generalising Herzberg's [31] results, we will have to assume that $\mathcal{C}$ is a compact non-empty convex subset of a given Banach space W (with norm $\|\cdot\|_W$).

3.1.2 Utilities

We fix some class $\mathcal{M}$ of functions from $\mathcal{C}$ to $\mathbb{R}$. The elements of $\mathcal{M}$ are called *admissible utility functions*. Every individual's utility function, u_i, belongs to $\mathcal{M}$. Elements of $\mathcal{M}^N$ will be called *utility profiles* and $\underline{u} = (u_i)_{i \in N} \in \mathcal{M}^N$.

3.1.3 Aggregation

We now employ some social choice theory notations and formulations from Kirman and Sondermann [35].

Definition 3.1. A relation $P \subseteq \mathcal{C} \times \mathcal{C}$ is called a *weak order* if and only if P satisfies the following properties:

- P is *asymmetric*: $\forall x, y \in \mathcal{C} \quad (xPy \Rightarrow \neg yPx)$.
- P is *negatively transitive*: $\forall x, y, z \in \mathcal{C} \quad (\neg xPy \wedge \neg yPz \Rightarrow \neg xPz)$.

Lemma 3.2. (Theorem 2.1. in Fishburn [22]). *A weak order $P \subseteq \mathcal{C} \times \mathcal{C}$ is transitive.*

Proof. Since P is negatively transitive, for all $x, y, z \in \mathcal{C}$ we have

$$xPy \Rightarrow (xPz \vee zPy). \tag{3.1}$$

Indeed, suppose that (3.1) is not true, i.e., $(xPy \wedge \neg xPz \wedge \neg zPy)$. By negative transitivity, $\neg xPz$ and $\neg zPy$ implies $\neg xPy$, which contradicts xPy. Therefore, negative transitivity implies (3.1). Now suppose xPy and yPz. Then by (3.1) we have

$$(xPz \vee zPy) \quad \wedge \quad (yPx \vee xPz).$$

Since zPy and yPx are false by asymmetry, xPz is true. Thus P is transitive. □

Remark 3.3. Due to the previous lemma, a weak order P on $\mathcal{C}$ is transitive and satisfies

$$((\neg yPx \wedge yPz) \quad \vee \quad (xPy \wedge \neg zPy)) \Rightarrow xPz, \tag{3.2}$$

for all $x, y, z \in \mathcal{C}$.

Notation 3.4. *For all $x, y \in \mathcal{C}$, when we write xPy it means that $(x, y) \in P$ and should be read as 'x is preferred to y'.*

Definition 3.5. $\mathcal{P}$ denotes the set of all weak orders[1] on $\mathcal{C}$. For all $x, y \in \mathcal{C}$ and $\underline{P} = (P_i)_{i \in N} \in \mathcal{P}^N$, we define the *coalition supporting x over y under $\underline{P}$* as follows:

$$C(x, y, \underline{P}) := \{i \in N \, : \, xP_i y\}.$$

Definition 3.6. For $x \in \mathcal{C}$ and $P \in \mathcal{P}$, x will be called *P-maximal* if and only if for all $y \in \mathcal{C} \backslash \{x\}$, we have xPy.

[1] Do not confuse this notation with the power-set notation.

Definition 3.7. For $u : \mathcal{C} \longrightarrow \mathbb{R}$ and $P \in \mathcal{P}$, we say that u is a *utility representation* of P if and only if for all $x, y \in \mathcal{C}$,

$$u(x) > u(y) \Leftrightarrow xPy.$$

Notation 3.8. *We use $P^u \in \mathcal{P}$ to denote that the utility function u, induces the preference P. Similarly, given an N-sequence $\underline{u} = (u_i)_{i \in N}$ of functions from $\mathcal{C}$ to $\mathbb{R}$, we define*

$$\underline{P}^{\underline{u}} := (P^{u_i})_{i \in N} \in \mathcal{P}^N.$$

We say that the utility profile $\underline{u}$ induces the preference profile $\underline{P}^{\underline{u}}$.

Definition 3.9. A *social welfare function (with universal domain)* is a map

$$\sigma : \mathcal{P}^N \longrightarrow \mathcal{P}.$$

According to our notations, Arrow's rationality axioms for σ will be formulated as follows:

Axiom 3.10. (Unanimity Preservation). *For all $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$, if $C(x, y, \underline{P}) = N$ then $x\sigma(\underline{P})y$.*

Axiom 3.11. (Independence of Irrelevant Alternatives). *For all $x, y \in \mathcal{C}$ and $\underline{P}, \underline{P}' \in \mathcal{P}^N$, if $C(x, y, \underline{P}) = C(x, y, \underline{P}')$ and $C(y, x, \underline{P}) = C(y, x, \underline{P}')$, then*

$$x\sigma(\underline{P})y \Leftrightarrow x\sigma(\underline{P}')y, \qquad y\sigma(\underline{P})x \Leftrightarrow y\sigma(\underline{P}')x.$$

Axiom 3.12. (No Dictatorship). *There is no $i_0 \in N$ such that for all $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$,*

$$xP_{i_0}y \Rightarrow x\sigma(\underline{P})y.$$

Definition 3.13. We say that a coalition $C \subseteq N$ is *σ-decisive* if and only if for all $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$ one has $x\sigma(\underline{P})y$ whenever xP_iy for all $i \in C$. The set of σ-decisive coalitions is denoted by F_σ.

Remark 3.14. In notation, a coalition $C \subseteq N$ is σ-decisive if and only if for all $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$ one has

$$(\forall i \in C \quad (xP_iy)) \Rightarrow x\sigma(\underline{P})y.$$

3.2 Kirman-Sondermann correspondence

For the following, recall from Section 2, that an *ultrafilter* on N is a filter F on N which is maximal with respect to inclusion. We have shown in Section 2.1.3 that a filter F is an ultrafilter if and only if for all $A \subseteq N$, either $A \in F$ or $N \backslash A \in F$. An ultrafilter is *non-principal* if the intersection of all its members is empty. Otherwise it is called *principal*, and one can show that the intersection has exactly one element (in our interpretation a *dictator*).

Let Σ be the set of social welfare functions satisfying the Arrovian rationality axioms (Axioms 3.10 and 3.11).

Definition 3.15. For any $\sigma \in \Sigma$, define the following subsets of the power-set of N:

$$F := \left\{C \subseteq N \,:\, \exists x, y \in \mathcal{C}\ \exists \underline{P} \in \mathcal{P}^N\ \ (\forall i \in C\ \forall j \in N\backslash C \quad xP_iy \wedge yP_jx \wedge x\sigma(\underline{P})y)\right\},$$
$$F' := \left\{C \subseteq N \,:\, \exists x, y \in \mathcal{C}\ \forall \underline{P} \in \mathcal{P}^N\ \ (\forall i \in C\ \forall j \in N\backslash C \quad (xP_iy \wedge yP_jx)) \Rightarrow x\sigma(\underline{P})y\right\},$$
$$F'' := \left\{C \subseteq N \,:\, \forall x, y \in \mathcal{C}\ \forall \underline{P} \in \mathcal{P}^N\ \ (\forall i \in C\ \forall j \in N\backslash C \quad (xP_iy \wedge yP_jx)) \Rightarrow x\sigma(\underline{P})y\right\}.$$

Lemma 3.16. (Kirman and Sondermann [35]). *Let F, F', F'' be as above, then*

1. $F = F' = F''$, *and*
2. F *is an ultrafilter.*

Proof. 1. We know that $F'' \subset F' \subset F$. We have to show that $F \subset F' \subset F''$. From Axiom 3.11 $F \subset F'$. It only remains to prove that $F' \subset F''$. Take $C \in F'$ and $z \in \mathcal{C}$ with $z \neq x$ and $z \neq y$. Now assume that there exists $\underline{P} \in \mathcal{P}^N$ such that xP_iz for all $i \in C$, and zP_jx for all $j \in N\backslash C$. Consider $\underline{P}' \in \mathcal{P}^N$, such that

$$\forall i \in C \quad xP'_iy \wedge yP'_iz, \tag{3.3}$$
$$\forall j \in N\backslash C \quad yP'_jz \wedge zP'_jx. \tag{3.4}$$

By transitivity of $\underline{P}'$, we have

$$\forall i \in C \quad \forall j \in N\backslash C \qquad xP'_iz \wedge yP'_jx. \tag{3.5}$$

Since $C \in F'$,

$$\big(\forall i \in C \quad \forall j \in N\backslash C \quad \big(xP'_iy \wedge yP'_jx\big)\big) \Rightarrow x\sigma(\underline{P}')y.$$

By (3.3) and (3.4), $C(y, z, \underline{P}') = N$, and thus by Axiom 3.10, $y\sigma(\underline{P}')z$. Therefore by transitivity, $x\sigma(\underline{P}')z$. However from (3.4) and (3.5), $C(x, z, \underline{P}) = C(x, z, \underline{P}')$ and $C(z, x, \underline{P}) = C(z, x, \underline{P}')$. Thus by Axiom 3.11 $x\sigma(\underline{P})z$. With a similar argument we get

$$(\forall i \in C \quad \forall j \in N\backslash C \quad (zP_iy \wedge yP_jz)) \Rightarrow z\sigma(\underline{P})y. \tag{3.6}$$

Now take $w \in \mathcal{C}$ with $w \neq x$, $w \neq y$ and $w \neq z$. If we replace x by w in the above argument, for all $\underline{P} \in \mathcal{P}^N$ we have

$$(\forall i \in C \quad \forall j \in N \backslash C \quad (wP_i z \wedge zP_j w)) \Rightarrow w\sigma(\underline{P})z.$$

Finally, if we replace w by y in (3.6), for all $\underline{P} \in \mathcal{P}^N$ we have

$$(\forall i \in C \quad \forall j \in N \backslash C \quad (zP_i w \wedge wP_j z)) \Rightarrow z\sigma(\underline{P})w,$$

and this completes the proof of the first part.

2. We have to verify the ultrafilter properties[2] for F. First, since σ satisfies Axiom 3.10, the empty set cannot belong to F. The next property which we have to verify is that F is closed under intersections. Let $C_1, C_2 \in F$. Put

$$\begin{aligned} D_1 &:= C_1 \cap C_2, \\ D_2 &:= C_1 \cap (N \backslash C_2), \\ D_3 &:= C_2 \cap (N \backslash C_1), \\ D_4 &:= N \backslash (C_1 \cup C_2). \end{aligned}$$

Choose $x, y, z \in \mathcal{C}$. We define $\underline{P} \in \mathcal{P}^N$ as follows:

$$\begin{aligned} &\forall i \in D_1 \qquad zP_i x \wedge xP_i y, \\ &\forall i \in D_2 \qquad xP_i y \wedge yP_i z, \\ &\forall i \in D_3 \qquad yP_i z \wedge zP_i x, \\ &\forall i \in D_4 \qquad yP_i x \wedge xP_i z. \end{aligned}$$

From $C_1 = D_1 \cup D_2 \in F = F''$ it follows that $x\sigma(\underline{P})y$. Also $C_2 = D_1 \cup D_3 \in F = F''$ implies that $z\sigma(\underline{P})x$. Thus by transitivity of $\sigma(\underline{P})$, $z\sigma(\underline{P})y$, which proves

$$D_1 = C_1 \cap C_2 \in F.$$

Now we prove that if $C_1 \subseteq N$, then either $C_1 \in F$ or $N \backslash C_1 \in F$. Take an element $a \in \mathcal{C}$ and a weak order P on $\mathcal{C} \backslash \{a\}$. Define $\underline{P} \in \mathcal{P}^N$ as follows:

$$\begin{aligned} &\forall i \in N && P_i = P, \\ &\forall i \in C_1 \quad \forall x \in \mathcal{C} \backslash \{a\} && xP_i a, \\ &\forall i \in N \backslash C_1 \quad \forall x \in \mathcal{C} \backslash \{a\} && aP_i x. \end{aligned} \tag{3.7}$$

[2] The ultrafilter properties are (i) $\emptyset \notin F$, (ii) if $C_1, C_2 \in F$, then $C_1 \cap C_2 \in F$, (iii) if $C_1 \in F$ and $C_1 \subseteq C_2$, then $C_2 \in F$, and (iv) if $C_1 \subseteq N$, then either $C_1 \in F$ or $N \backslash C_1 \in F$.

Since P is a weak order on $\mathcal{C}\backslash\{a\}$, we can assume that for $x, y \in \mathcal{C}\backslash\{a\}$, xPy. Hence by (3.7) and Axiom 3.10, $x\sigma(\underline{P})y$. By formula (3.1), $x\sigma(\underline{P})a$ or $a\sigma(\underline{P})x$. Therefore either $C_1 = \{i \in N : xP_i a\} \in F$ or $N\backslash C_1 = \{i \in N : aP_i x\} \in F$.

Finally we need to show that F is closed under supersets. Let $C_1 \in F$ and $C_1 \subseteq C_2 \subseteq N$. We have just proved that either $C_2 \in F$ or $N\backslash C_2 \in F$. But if $N\backslash C_2 \in F$, we have

$$\emptyset = (N\backslash C_2) \cap C_1 \in F,$$

which is impossible according to the property that empty set cannot belong to F. Hence F is an ultrafilter.

□

We have the following important theorem from Kirman and Sondermann [35]:

Theorem 3.17. (Kirman-Sondermann Correspondence). *Let Σ, as before, be the set of social welfare functions satisfying Arrovian rationality axioms (Axioms 3.10 and 3.11) and $\mathcal{F}$ be the set of all ultrafilters on N. Then*

1. *There is a bijection between Σ and $\mathcal{F}$, given by $\lambda : \Sigma \to \mathcal{F}$, $\sigma \mapsto F_\sigma$, where F_σ is as in Definition 3.13.*

2. *In addition, $\sigma \in \Sigma$ satisfies Axiom 3.12 if and only if the corresponding ultrafilter $\lambda(\sigma) = F_\sigma$ is non-principal.*

Proof. 1. According to the previous lemma, for all $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$, the set

$$F_\sigma = \{C \subseteq N : (\forall i \in C \quad \forall j \in N\backslash C \quad (xP_i y \wedge yP_j x)) \Rightarrow x\sigma(\underline{P})y\},$$

is an ultrafilter on N. Due to Definition 3.13, we have to show that if $C \in F_\sigma$ and $xP_i y$ for all $i \in C$, then $x\sigma(\underline{P})y$. For given $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$, put

$$\begin{aligned} D_1 &:= \{i \in N : xP_i y\}, \\ D_2 &:= \{i \in N : yP_i x\}, \\ D_3 &:= N\backslash(D_1 \cup D_2). \end{aligned}$$

Define $\underline{P}' = (P'_i)_{i \in N} \in \mathcal{P}^N$ for $x, y, z \in \mathcal{C}$ as follows:

$$\forall i \in D_1 \qquad xP'_i z \wedge zP'_i y, \tag{3.8}$$

$$\forall i \in D_2 \qquad yP'_i z \wedge zP'_i x, \tag{3.9}$$

$$\forall i \in D_3 \qquad \neg xP'_i y \wedge \neg yP'_i x \wedge xP'_i z. \tag{3.10}$$

By (3.2) and (3.10), we deduce $yP'_i z$ for all $i \in D_3$. Furthermore $C(x, y, \underline{P}) = C(x, y, \underline{P}')$ and $C(y, x, \underline{P}) = C(y, x, \underline{P}')$. Since $C \in F_\sigma$, $C \subseteq D_1$ and F_σ is closed under supersets, $D_1 \in F_\sigma$. Therefore for all $i \in D_1$ and $j \in N \backslash D_1$, $zP'_i y$ and $yP'_j z$ imply $z\sigma(\underline{P}')y$. Again, since F_σ is closed under supersets, $C' = D_1 \cup D_3 \in F_\sigma$. Therefore

$$\big(\forall i \in C' \quad \forall j \in N \backslash C' \quad \big(xP'_i z \wedge zP'_j x\big)\big) \Rightarrow x\sigma(\underline{P}')z.$$

By transitivity of $\sigma(\underline{P}')$, $x\sigma(\underline{P}')z$ and $z\sigma(\underline{P}')y$ imply $x\sigma(\underline{P}')y$. Since we have $C(x, y, \underline{P}) = C(x, y, \underline{P}')$ and $C(y, x, \underline{P}) = C(y, x, \underline{P}')$, by Axiom 3.11, $x\sigma(\underline{P}')y$ if and only if $x\sigma(\underline{P})y$. Hence we showed that $\lambda(\sigma) = F_\sigma$ is well-defined. Suppose that there exists an $F'_\sigma \in \mathcal{F}$ with the same properties as F_σ, such that $F'_\sigma \neq F_\sigma$. There exists an element $C \in F'_\sigma$ which does not belong to F_σ. Since F_σ is an ultrafilter, $N \backslash C \in F_\sigma$, which is a contradiction.

Now, we prove that the mapping $\sigma \mapsto F_\sigma$ is surjective. Given an $F \in \mathcal{F}$, we have to show that there exists a $\sigma_0 \in \Sigma$ such that $\lambda(\sigma_0) = F$. For arbitrary $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$, we define σ_0 as follows:

$$x\sigma_0(\underline{P})y \Leftrightarrow C(x, y, \underline{P}) \in F.$$

σ_0 satisfies Axiom 3.10, since if $C(x, y, \underline{P}) = N$, we know that $N \in F$, and hence by the definition of σ_0, $x\sigma_0(\underline{P})y$. σ_0 also satisfies Axiom 3.11, since for all $\underline{P}, \underline{P}' \in \mathcal{P}^N$ and $x, y \in \mathcal{C}$, $x\sigma_0(\underline{P})y$ if and only if $C(x, y, \underline{P}) \in F$. If $C(x, y, \underline{P}) = C(x, y, \underline{P}')$, then also $C(x, y, \underline{P}') \in F$, and this is true if and only if $x\sigma_0(\underline{P}')y$. We have to show that for all $\underline{P} \in \mathcal{P}^N$, $\sigma_0(\underline{P}) \in \mathcal{P}$, i.e., is a weak order. Since $C(x, y, \underline{P}) \in F$ implies $C(y, x, \underline{P}) \notin F$, $\sigma_0(\underline{P})$ is asymmetric. Now assume that for all $x, y, z \in \mathcal{C}$,

$$\neg x\sigma_0(\underline{P})y \wedge y\sigma_0(\underline{P})z.$$

This means that $C_1 := C(x, y, \underline{P}) \notin F$, $C_2 := C(y, z, \underline{P}) \notin F$ and $C_1 \cup C_2 \notin F$[3]. Formula (3.1) implies

$$C_3 := C(x, z, \underline{P}) \subset C_1 \cup C_2,$$

hence $C_3 \notin F$, which proves that $\sigma_0(\underline{P})$ is negative transitive. Thus we have shown that $\sigma_0 \in \Sigma$. It only remains to show that the ultrafilter $F_{\sigma_0} = \lambda(\sigma_0)$ is equal to F. It suffices to prove that $F_{\sigma_0} \subset F$. For arbitrary $C_0 \in F_{\sigma_0}$, there exist $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$ such that $C_0 = C(x, y, \underline{P})$. Due to the properties of F_{σ_0}, this implies that $x\sigma_0(\underline{P})y$, which by the definition of σ_0, yields $C_0 \in F$.

2. First we show that if a $\sigma \in \Sigma$ satisfies Axiom 3.12, then the corresponding ultrafilter $\lambda(\sigma) = F_\sigma$ is non-principal. Suppose that F_σ is a principal ultrafilter on N and we know that this kind of ultrafilter is of the form $\{C \subseteq N : i_0 \in C\}$ for some $i_0 \in N$. Thus by the first part of the theorem, for all $C \in F_\sigma$, $x, y \in \mathcal{C}$ and

[3] An ultrafilter satisfies a contraction principle, i.e., $C_1 \cup C_2 \in F$ implies either $C_1 \in F$ or $C_2 \in F$.

$\underline{P} \in \mathcal{P}^N$ we have

$$(\forall i \in C \quad (xP_iy)) \Rightarrow x\sigma(\underline{P})y.$$

Since this is true for every element in C, also for $i_0 \in C$, $xP_{i_0}y \Rightarrow x\sigma(\underline{P})y$. Hence i_0 is a dictator.

Conversely, suppose that a $\sigma \in \Sigma$ is dictated by a $i_0 \in N$. Then $\{i_0\} \in F_\sigma$, which implies $\bigcap_{A \in F_\sigma} A \neq \emptyset$. For otherwise, $\{i_0\} \notin F_\sigma$, and since F_σ is an ultrafilter, $N \backslash \{i_0\} \in F_\sigma$. Thus the set of individuals is decisive against i_0.

□

3.3 Assumptions

In the previous section, we have shown that a social welfare function σ satisfies all Arrovian rationality axioms (Axioms 3.10, 3.11 and 3.12) if and only if the corresponding ultrafilter F_σ is a non-principal ultrafilter. We therefore require that $\mathcal{D}$, which is defined in Section 3.1, is a non-principal ultrafilter on N and this is only possible if N is infinite[4]. Thus, we impose the following assumption:

Assumption 3.18. *$\mathcal{D}$ is a non-principal ultrafilter on N (and therefore N is infinite).*

Parametrisations are ubiquitous in macroeconomics, motivating our next assumption:

Assumption 3.19. *Let Z be a compact subset of a given Banach space X (with norm $\|\cdot\|_X$). There exists a continuous function $v : Z \times \mathcal{C} \to \mathbb{R}$ such that for every $z \in Z$, $v(z, \cdot)$ is strictly concave and* [5]

$$\mathcal{M} \subseteq \{\, v(z, \cdot) \,:\, z \in Z \,\}.$$

(In other words, given any $\underline{u} \in \mathcal{M}^N$, there is an N-sequence $(z_i)_{i \in N} \in (Z)^N$ such that $u_i = v(z_i, \cdot)$ for every $i \in N$.)

Remark 3.20. For all $u \in \mathcal{M}$, u attains its unique global maximum on $\mathcal{C}$.

Proof. We are concerned with the maximisation problem

$$\forall z \in Z \qquad \max_{r \in \mathcal{C}} v(z, r). \tag{3.11}$$

Since $Z \times \mathcal{C}$ is compact and v is continuous, $v(Z \times \mathcal{C})$ is a compact (closed and bounded) subset of $\mathbb{R}$. According to boundedness of $v(Z \times \mathcal{C})$ in $\mathbb{R}$, it has a least upper bound α and since $v(Z \times \mathcal{C})$ is closed, it contains α. Therefore problem (3.11) has a solution.

[4]If N is finite, then every ultrafilter on it is principal.
[5]Properness is obvious, since v takes real values.

For proving the uniqueness we suppose that for problem (3.11) there exist two different solutions r_1 and r_2 in $\mathcal{C}$. Then from the convexity of $\mathcal{C}$, $\frac{r_1+r_2}{2}$ is also a solution. But $v(z,\cdot)$ is strictly concave for every $z \in Z$ and we have

$$v\left(z, \frac{r_1+r_2}{2}\right) > \frac{1}{2}\left(v(z,r_1)+v(z,r_2)\right) = \alpha,$$

which is a contradiction. □

3.4 Socially acceptable and representative utility functions

In this section we will show the existence of $\mathcal{D}$-socially acceptable and representative utility functions.

Definition 3.21. An admissible utility function $\varphi : \mathcal{C} \to \mathbb{R}$ is said to be $\mathcal{D}$-*socially acceptable* for $\underline{u}$ if and only if there exists some $\tilde{x} \in \mathcal{C}$ with $\varphi(\tilde{x})=\sup \varphi$ such that for every $y \in \mathcal{C}\backslash\{\tilde{x}\}$, the coalition of i with $u_i(\tilde{x}) > u_i(y)$ is decisive.

Definition 3.22. An admissible utility function $\varphi : \mathcal{C} \to \mathbb{R}$ is called σ-*representative* of $\underline{P} \in \mathcal{P}^N$ if and only if there exists some $\tilde{x} \in \mathcal{C}$ with $\varphi(\tilde{x})=\sup \varphi$ and any such $\tilde{x}$ is also $\sigma(\underline{P})$-maximal.

Theorem 3.23. *Suppose Assumptions 3.18 and 3.19 hold. There exists for every* $\underline{u} \in \mathcal{M}^N$ *some* $\mathcal{D}$*-socially acceptable utility function.*

Proof. Fix an arbitrary $\underline{u} \in \mathcal{M}^N$ and by Assumption 3.19, let $(z_i)_{i\in N} \in (Z)^N$ be such that $u_i = v(z_i,\cdot)$ for every $i \in N$. The ultrapower construction can easily be adapted to construct an embedding

$$^* : V\left((X \oplus W) \cup \mathbb{R}\right) \to V(^*(X \oplus W) \cup {}^*\mathbb{R}),$$

where $^*(X \oplus W)$ is a Banach space over $^*\mathbb{R}$ satisfying:

- *extension*: $^*(X \oplus W)$ and $^*\mathbb{R}$ are extensions of $X \oplus W$ and $\mathbb{R}$, respectively, and $^*x = x$ for all $x \in (X \oplus W) \cup \mathbb{R}$, and
- *transfer*: If $\Phi(v_1, \cdots, v_n)$ is an $\in$-formula with bounded quantifiers and n free variables, then for all $A_1, \cdots, A_n \in V\left((X \oplus W) \cup \mathbb{R}\right)$,

$$V\left((X \oplus W) \cup \mathbb{R}\right) \models \Phi[A_1, \cdots, A_n] \quad \Leftrightarrow \quad V(^*(X \oplus W) \cup {}^*\mathbb{R}) \models \Phi[^*A_1, \cdots, {}^*A_n].$$

For the rest of the proof, we work in the resulting nonstandard universe. We have to construct some parameter $\tilde{z}$ such that $v(\tilde{z},\cdot)$ is $\mathcal{D}$-socially acceptable. Let

$$\bar{z} := [(z_i)_{i\in N}]_{\mathcal{D}} \in {}^*Z.$$

Z is compact, then by Proposition 2.87, every element of ${}^*Z$ is nearstandard and let $\tilde{z} := {}^\circ\bar{z}$. Applying the Transfer Principle of nonstandard analysis to Remark 3.20, we learn that ${}^*v(\bar{z},\cdot)$ attains its unique global *maximum in some $\bar{x} \in {}^*\mathcal{C}$.

Consider now the map

$$w : Z \to \mathcal{C},$$

which assigns to each $z \in Z$ the unique $x = w(z) \in \mathcal{C}$ such that (existence and uniqueness follow from Remark 3.20)

$$x \in \arg\sup_{r\in\mathcal{C}} v(z,r).$$

By the Transfer Principle,

$${}^*w : {}^*Z \to {}^*\mathcal{C},$$

hence ${}^*w(\bar{z}) \in {}^*\mathcal{C}$ and since $\mathcal{C}$ is a compact, every element of ${}^*\mathcal{C}$ is nearstandard (see again Proposition 2.87) and therefore $\bar{x} = {}^*w(\bar{z})$ is nearstandard. We put $\tilde{x} := {}^\circ\bar{x}$.

Due to Assumption 3.19, v is continuous and hence by Proposition 2.89, ${}^*v$ is S-continuous. Therefore, we have for all $y \in \mathcal{C}$,

$$v(\tilde{z},y) - v(\tilde{z},\tilde{x}) \approx {}^*v(\bar{z},y) - {}^*v(\bar{z},\tilde{x}) \approx {}^*v(\bar{z},y) - {}^*v(\bar{z},\bar{x}). \tag{3.12}$$

The right-hand side of equation (3.12) is a non-positive hyperreal (since $\bar{x}$ is a global *maximum of ${}^*v(\bar{z},\cdot)$), so the standard part is non-positive, but the standard part is exactly $v(\tilde{z},y) - v(\tilde{z},\tilde{x})$; then

$$v(\tilde{z},y) - v(\tilde{z},\tilde{x}) \leq 0 \quad ; \qquad \forall y \in \mathcal{C}.$$

Since we have a unique global maximum (according to Remark 3.20),

$$v(\tilde{z},y) - v(\tilde{z},\tilde{x}) < 0 \quad ; \qquad \text{for all}\, y \neq \tilde{x}. \tag{3.13}$$

The rest of the proof is basically as in Herzberg [31]. In order to verify that $v(\tilde{z},\cdot)$ is $\mathcal{D}$-socially acceptable, we still need to show that for every $y \in \mathcal{C}\backslash\{\tilde{x}\}$, the set of all $i \in N$ with $u_i(\tilde{x}) > u_i(y)$ is decisive (i.e. $\in \mathcal{D}$). Define a function f by $f(h) := v(h,\tilde{x}) - v(h,y)$ for all $h \in Z$, whence

$$\{i \in N \,:\, u_i(\tilde{x}) > u_i(y)\,\} = \{i \in N \,:\, f(z_i) > 0\,\}. \tag{3.14}$$

Due to the construction of the nonstandard embedding * via the bounded ultrapower (with respect to $\mathcal{D}$) of the superstructure $V\left((X \oplus W) \cup \mathbb{R}\right)$, one has the equivalence[6]

$$\{i \in N \,:\, u_i(\tilde{x}) > u_i(y)\,\} \in \mathcal{D} \Leftrightarrow {}^*f(\bar{z}) > 0. \tag{3.15}$$

[6]The sequence $(f(z_i))_{i\in N}$ is bounded in $V\left((X \oplus W) \cup \mathbb{R}\right)$, since $z_i \in Z \subseteq X$ for all $i \in N$ and $f : Z \to \mathbb{R}$.

However, by applying the Transfer Principle to the defining equation for f and due to S-continuity of ${}^*v$, we get

$$ {}^*f(\bar{z}) = {}^*v(\bar{z},\tilde{x}) - {}^*v(\bar{z},y) \approx {}^*v(\bar{z},\bar{x}) - {}^*v(\bar{z},y). $$

The standard part of the right-hand side is strictly positive (by inequality (3.13)) and therefore ${}^\circ({}^*f(\bar{z})) > 0$. Hence ${}^*f(\bar{z}) > 0$ and by equivalence (3.15) we have

$$ \{ i \in N \,:\, u_i(\tilde{x}) > u_i(y) \} \in \mathcal{D}. $$

□

Theorem 3.24. *Suppose σ satisfies Axioms 3.10, 3.11 and 3.12. Then:*

1. *$\mathcal{D} := F_\sigma$ satisfies Assumption 3.18.*
2. *If, in addition, $\mathcal{M}$ satisfies Assumption 3.19, then there exists for every $\underline{u} \in \mathcal{M}^N$ some admissible utility function which is σ-representative of the preference profile $\underline{P^u}$ induced by $\underline{u}$.*

Proof. The proof follows the same lines as the proof of Theorem 2 in Herzberg [31].

1. We have shown in Theorem 3.17 that $\mathcal{D} = F_\sigma$ is a non-principal ultrafilter whenever σ satisfies Axioms 3.10, 3.11 and 3.12.
2. According to Theorem 3.23, for an arbitrary $\underline{u} \in \mathcal{M}^N$ there exist some $\varphi \in \mathcal{M}$ and $\tilde{x} \in \mathcal{C}$ such that $\varphi(\tilde{x})=$ sup φ and for every $y \in \mathcal{C}\backslash\{\tilde{x}\}$

$$ \{ i \in N \,:\, u_i(\tilde{x}) > u_i(y) \} \in F_\sigma. $$

Now fix an arbitrary $y \in \mathcal{C}\backslash\{\tilde{x}\}$. For all $i \in N$ with $u_i(\tilde{x}) > u_i(y)$ we have $\tilde{x}P^{u_i}y$ and thus $\tilde{x}P_i^{\underline{u}}y$ by definition. Therefore

$$ \{ i \in N \,:\, \tilde{x}P_i^{\underline{u}}y \} \supseteq \{ i \in N \,:\, u_i(\tilde{x}) > u_i(y) \} \in F_\sigma. $$

Since F_σ is an ultrafilter,

$$ \{ i \in N \,:\, \tilde{x}P_i^{\underline{u}}y \} \in F_\sigma, $$

and this implies $\tilde{x}\sigma(\underline{P^u})y$ (part one in Theorem 3.17). The proof is complete since y was chosen arbitrarily in $\mathcal{C}\backslash\{\tilde{x}\}$.

□

Chapter 4

Illustration: Possible macroeconomic applications

In this chapter we provide sufficient conditions for the preceding theorems (Theorems 3.23 and 3.24 in Chapter 3) to be satisfied in economic applications. After this modification, we show a macroeconomic application as an example. We establish some model of the economy which consists of the government and the agents with their own preferences. Then we prove that under some assumptions, there exists a socially optimal path for the economy, which satisfies the government and the agents' maximisation problems. In the end, we consider some special cases for the maximisation.

4.1 Applicability of the preceding theorems

As before, suppose $\mathcal{C}$ is the set of social alternatives which is a compact non-empty convex subset of a given Banach space W. Consider a function u, defined by

$$\forall c \in \mathcal{C} \qquad u(c) = \max_{y:(y,c)\in\mathcal{Y}} f(y,c), \tag{4.1}$$

where $\mathcal{Y}$ is a set (called the choice set) and $f : \mathcal{Y} \to \mathbb{R}$ is some function, interpreted as a *happiness function* of some agent depending on the social parameter c and his (or her) choice y, where the optimal y, given c, depends on the social parameter c.

Lemma 4.1. *Given a Banach space B (with norm $\|\cdot\|_B$), suppose $\mathcal{Y}$ is a compact non-empty convex subset of $B \times \mathcal{C}$ and $f : \mathcal{Y} \to \mathbb{R}$ is continuous and strictly concave. Then u defined by* (4.1) *is continuous and strictly concave.*

Proof. (First proof). Since f is strictly concave, for every $(y', c'), (y'', c'') \in \mathcal{Y}$ and $\lambda \in (0, 1)$ we have

$$f\left(\lambda(y', c') + (1-\lambda)(y'', c'')\right) > \lambda f(y', c') + (1-\lambda) f(y'', c''). \tag{4.2}$$

Because $\mathcal{Y}$ is a compact set and f is continuous, $f(\mathcal{Y})$ is a compact (closed and bounded) subset of $\mathbb{R}$. According to boundedness of $f(\mathcal{Y})$ in $\mathbb{R}$, it has a least upper bound α and since $f(\mathcal{Y})$ is closed, it contains α. Therefore there exists a maximiser on $\mathcal{Y}$.

Now we suppose that there exist $y', y'' \in B$ such that

$$\forall c' \in \mathcal{C} \qquad f(y', c') = \max_{y:(y,c')\in\mathcal{Y}} f(y, c') = u(c'), \tag{4.3}$$

$$\forall c'' \in \mathcal{C} \qquad f(y'', c'') = \max_{y:(y,c'')\in\mathcal{Y}} f(y, c'') = u(c''). \tag{4.4}$$

For any $c', c'' \in \mathcal{C}$ and $\lambda \in (0, 1)$, we can write

$$\begin{aligned}
\lambda u(c') + (1-\lambda)u(c'') &= \lambda f(y', c') + (1-\lambda) f(y'', c'') && \text{(by (4.3) and (4.4))}\\
&< f\left(\lambda(y', c') + (1-\lambda)(y'', c'')\right) && \text{(by (4.2))}\\
&= f\left(\lambda y' + (1-\lambda)y'', \lambda c' + (1-\lambda)c''\right)\\
&\leq \max_{y:(y,\lambda c'+(1-\lambda)c'')\in\mathcal{Y}} f(y, \lambda c' + (1-\lambda)c'') && \text{(by convexity of } \mathcal{Y})\\
&= u\left(\lambda c' + (1-\lambda)c''\right), && \text{(by definition of } u)
\end{aligned}$$

and therefore u is strictly concave.

Let $c \in \mathcal{C}$ arbitrary and $(c_n)_n$ be a sequence converging to c. By definition of largest subsequential limit, $\limsup u(c_n)$, there exists $(c_{n_l})_l$ such that

$$\limsup_{n\to\infty} u(c_n) = \lim_{l\to\infty} u(c_{n_l}). \tag{4.5}$$

For all $n \in \mathbb{N}$, we define

$$\mathcal{Y}_{c_n} := \{\, y \in B \quad : \quad (y, c_n) \in \mathcal{Y} \,\},$$

which is a compact set[1].

Since $\mathcal{Y}_{c_n}$ is compact and $f(\cdot, c_n)$ is continuous,

$$\exists y_n \in \arg\max_{\mathcal{Y}_{c_n}} f(\cdot, c_n).$$

[1]For proving the compactness of $\mathcal{Y}_{c_n}$, let $(y_k)_k$ be a sequence in $\mathcal{Y}_{c_n}$ such that for every $k \in \mathbb{N}$, $(y_k, c_n) \in \mathcal{Y}$. Since $\mathcal{Y}$ is compact, there exists $(y_{k_l})_l$ in B such that

$$\lim_{l\to\infty} (y_{k_l}, c_n) = (y, c_n) \in \mathcal{Y}.$$

By definition of $\mathcal{Y}_{c_n}$, it means that $y \in \mathcal{Y}_{c_n}$. Therefore from any sequence in $\mathcal{Y}_{c_n}$ we can extract a subsequence with a limit in $\mathcal{Y}_{c_n}$ and this implies the compactness of $\mathcal{Y}_{c_n}$ (see Characterisation of Compactness Theorem in Royden and Fitzpatrick [54]).

By compactness of $\mathcal{Y}$,

$$\exists (y_{n_{l_k}})_k \quad \exists y \in B \qquad \text{s.t.} \qquad y_{n_{l_k}} \xrightarrow{k\to\infty} y,$$

therefore

$$(y_{n_{l_k}}, c_{n_{l_k}}) \xrightarrow{k\to\infty} (y, c). \tag{4.6}$$

Let us suppose that $y \in \arg\max f(\cdot, c)$, which will be proved later. Then

$$\begin{aligned}
\limsup_{n\to\infty} u(c_n) &= \lim_{k\to\infty} u(c_{n_{l_k}}) && \text{(by (4.5))}\\
&= \lim_{k\to\infty} f(y_{n_{l_k}}, c_{n_{l_k}}) && \text{(since } y_{n_{l_k}} \in \arg\max_{\mathcal{Y}_{c_{n_{l_k}}}} f(\cdot, c_{n_{l_k}}))\\
&= f(y, c) && \text{(by continuity of } f \text{ and (4.6))}\\
&= u(c). && \text{(since } y \in \arg\max f(\cdot, c))
\end{aligned}$$

Similarly it can be shown that $\liminf_{n\to\infty} u(c_n) = u(c)$. Hence

$$\lim_{n\to\infty} u(c_n) = u(c),$$

which proves the continuity of u.

Suppose that $y \notin \arg\max f(\cdot, c)$, i.e. there exists a $\hat{y}$ with $(\hat{y}, c) \in \mathcal{Y}$ such that

$$f(\hat{y}, c) > f(y, c).$$

By the continuity of f, we get

$$\lim_{k\to\infty} f(\hat{y}, c_{n_{l_k}}) = f(\hat{y}, c) > f(y, c) = \lim_{k\to\infty} f(y_{n_{l_k}}, c_{n_{l_k}}).$$

This implies that for sufficiently large k,

$$f(\hat{y}, c_{n_{l_k}}) > f(y_{n_{l_k}}, c_{n_{l_k}}),$$

which means $(y_{n_{l_k}}, c_{n_l})$ is not a maximiser and this is a contradiction to

$$y_{n_{l_k}} \in \arg\max f(\cdot, c_{n_{l_k}}),$$

and this completes the proof of the lemma. □

Proof. (Second proof). u is strictly concave by the first proof. Since f is a continuous function on the compact set $\mathcal{Y}$, f is uniformly continuous (see Royden and Fitzpatrick [54] [Chapter 9; Proposition 23]). This means that for every $\epsilon > 0$ there exists a $\delta > 0$

such that

$$\forall (y', c'), (y'', c'') \in \mathcal{Y} \qquad \| (y', c') - (y'', c'')\|_\infty < \delta \implies \big| f(y', c') - f(y'', c'') \big| < \epsilon.$$

In other words, for any $\epsilon > 0$ there is a $\delta > 0$ with

$$f(y', c') - \epsilon < f(y'', c'') < f(y', c') + \epsilon, \tag{4.7}$$

whenever

$$\| y' - y''\|_B < \delta \qquad \text{and} \qquad \| c' - c''\|_W < \delta.$$

For fixed $c'' \in \mathcal{C}$, there is a point[2] $y'' \in B$ with $u(c'') = f(y'', c'')$. Then

$$\begin{aligned} f(y', c') - \epsilon &< u(c'') \\ &< f(y'', c') + \epsilon && \text{(by (4.7))} \\ &\leq \max_{y:(y,c') \in \mathcal{Y}} f(y, c') + \epsilon \\ &= u(c') + \epsilon, && \text{(by (4.3))} \end{aligned}$$

for all c' in $\mathcal{C}$ with $\| c' - c''\|_W < \delta$. By symmetry between c' and c'' it follows that $|u(c') - u(c'')| < \epsilon$ such that $\| c' - c''\|_W < \delta$, which proves that the function u is continuous on $\mathcal{C}$. □

Corollary 4.2. *Suppose $\mathcal{Y}$ is a compact non-empty convex subset of $B \times \mathcal{C}$ and $g : Z \times \mathcal{Y} \to \mathbb{R}$ is continuous, where Z is as in Assumption 3.19. Furthermore suppose that $g(z, \cdot)$ is strictly concave for all $z \in Z$. Let*

$$\mathcal{M} \subseteq \{ \max_{y:(y,\cdot) \in \mathcal{Y}} g(z, y, \cdot) \quad : \quad z \in Z \},$$

and $\mathcal{D}$ be a non-principal ultrafilter on N. Then there exists for every $\underline{u} \in \mathcal{M}^N$ some $\mathcal{D}$-socially acceptable utility function.

Corollary 4.3. *Suppose σ satisfies Axioms 3.10, 3.11 and 3.12, and let $\mathcal{Y}$, g and $\mathcal{M}$ be as in Corollary 4.2. Then there exists for every $\underline{u} \in \mathcal{M}^N$ some admissible utility function which is σ-representative of the preference profile $\underline{P}^{\underline{u}}$ induced by $\underline{u}$.*

4.2 An example of a possible macroeconomic application

For applying our previous results, we give an example of a possible macroeconomic application.

[2] As we have argued in the beginning of the first proof, existence of the maximiser is obvious by compactness of $\mathcal{Y}$ and continuity of f.

4.2.1 The economy

For establishing our model, we consider there exist agents and the government in the economy who hold some asset stocks at time $t-1$ and these assets pay the interest out at time t. Note that agents and the government carry their budget constraints over at each time. We also suppose that they can choose their budget elements from some sets which all belong to the l^p space[3] on $\mathbb{R}^\infty$, for $p \in [1, \infty]$.

We employ some social choice theory notations from the previous chapter, and some notations for monetary theory from Walsh [62].

4.2.1.1 The government

We fix a set $\mathcal{C}$, chosen by the government, as a set of *social alternatives*. For satisfying our results, we assume that $\mathcal{C}$ is a compact non-empty convex subset of l^p. The government needs to generate revenue for financing its consumptions and debts. We assume that it generates revenue by printing money. We also suppose that the government should pay interest only for debts held by the agents.

We denote the government's purchases at time t by G_t, which is taken to be exogenous from X, where X is a compact subset of l^p; and its payment of interest on debts by $i_{t-1}B_{t-1}$, where i_{t-1} is the nominal rate of interest between periods $t-1$ and t. These two expenditures can be financed with two sources. One by borrowing from the agents, $B_t - B_{t-1}$, and second by printing money, $M_t - M_{t-1}$. Therefore we can write the government's budget constraint[4] as following:

$$G_t + i_{t-1}B_{t-1} = (B_t - B_{t-1}) + (M_t - M_{t-1}) \, .$$

In this budget constraint there is no revenue created by taxes other than inflation. The government's budget constraint can be deflated by the price level P_t and we denote all the deflated values with lower-case letters. Then we have

$$g_t + i_{t-1}\left(\frac{b_{t-1}}{1+\pi_t}\right) = \left(b_t - \frac{b_{t-1}}{1+\pi_t}\right) + \left(m_t - \frac{m_{t-1}}{1+\pi_t}\right),$$

where $b_{t-1} = B_{t-1}/P_{t-1}$ is real debt and $\pi_t = \frac{P_t - P_{t-1}}{P_{t-1}}$ is inflation rate. By subtracting b_{t-1} from both sides of the last equation, we can write

$$g_t + r_{t-1}b_{t-1} = (b_t - b_{t-1}) + \left(m_t - \frac{m_{t-1}}{1+\pi_t}\right), \tag{4.8}$$

---

[3]Recall that for $1 \le p < \infty$, $l^p = \{x = (x_t)_{t\ge 0} \in \mathbb{R}^\infty : \left(\sum_{t=0}^{\infty} |x_t|^p\right)^{\frac{1}{p}} < \infty\}$ with the norm $\|x\|_p = \left(\sum_{t=0}^{\infty} |x_t|^p\right)^{\frac{1}{p}}$ and for $p = \infty$, $l^\infty = \{x = (x_t)_{t\ge 0} \in \mathbb{R}^\infty : \sup_{t\ge 0} |x_t| < \infty\}$ with the norm $\|x\|_\infty = \sup_{t\ge 0} |x_t|$. Note that l^p space for $p \in [1, \infty]$ is a Banach space.

[4]Walsh in [62] has developed an equation for the government's budget constraint by combining the budget identities of the Treasury and the central bank. For more details see Walsh [62].

where $r_{t-1} = [(1+i_{t-1})/(1+\pi_t)] - 1$ is the real rate of interest[5] from $t-1$ to t.

The last term in (4.8) is the government's *seigniorage* revenue at time t and we denote it by s_t. By adding and subtracting m_{t-1}, the seigniorage can be written as

$$s_t = (m_t - m_{t-1}) + \left(\frac{\pi_t}{1+\pi_t}\right) m_{t-1}, \tag{4.9}$$

the sum of (i) agents' desired increase in real money holdings at time t and (ii) the inflation tax on real money holdings carried over from time $t-1$.

4.2.1.2 The agents

The economy is populated by an infinite number of agents and we denote their set by N. Every agent chooses a pair of bond and money holding paths from $\mathcal{A}$, which is a subset of $l^p \times l^p$. We assume that every agent in each period receives an exogenous consumption c_t from X. They also receive interest payments on government debt held at the start of period t, $(1+r_{t-1})b_{t-1}$, where b_{t-1} is the number of real bonds held at the start of this period. Every agent has real money holdings that are carried into period t from period $t-1$, $m_{t-1}/(1+\pi_t)$. Therefore we can write the budget constraint of the agent as below:

$$c_t + m_t + b_t = (1+r_{t-1})b_{t-1} + \frac{m_{t-1}}{(1+\pi_t)}. \tag{4.10}$$

4.2.2 Preferences

In this section we introduce the preferences of the agents and the government by using the notations from Section 3.1.

There exists a *weak order* relation P on $\mathcal{C} \times \mathcal{C}$ and we denote the set of all weak orders by $\mathcal{P}$. For all x and y in $\mathcal{C}$ when we have xPy, we say that *x is preferred to y*. Elements of $\mathcal{P}^N$ are called *profiles.* Each agent n has preferences that can be represented by a utility function $u^n : \mathcal{C} \longrightarrow \mathbb{R}$. We fix some class $\mathcal{M}$ of functions from $\mathcal{C}$ to $\mathbb{R}$ such that every agent's utility function belongs to $\mathcal{M}$.

We assume the government (social planner) to maximise social welfare. Thus, the preferences of the government are described by a *social welfare function*, $\sigma : \mathcal{P}^N \longrightarrow \mathcal{P}$, which satisfies all *Arrovian rationality axioms.* Recall from Section 3.1 that Arrovian rationality axioms are:

- (Unanimity Preservation). For all $x, y \in \mathcal{C}$ and $\underline{P} = (P^n)_{n\in N} \in \mathcal{P}^N$, if

  $$\{n \in N \,:\, xP^n y\} = N,$$

  then $x\sigma(\underline{P})y$.

[5] The relationship between real and nominal rates of interest is $1 + i_{t-1} = (1 + r_{t-1})(1 + \pi_t)$.

- (Independence of Irrelevant Alternatives). For all $x, y \in \mathcal{C}$ and $\underline{P}, \underline{P}' \in \mathcal{P}^N$, if

$$\{n \in N : xP^n y\} = \{n \in N : xP'^n y\}$$

and

$$\{n \in N : yP^n x\} = \{n \in N : yP'^n x\},$$

then

$$x\sigma(\underline{P})y \Leftrightarrow x\sigma(\underline{P}')y, \qquad y\sigma(\underline{P})x \Leftrightarrow y\sigma(\underline{P}')x.$$

- (No Dictatorship). There is no $n_0 \in N$ such that for all $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$,

$$xP^{n_0}y \Rightarrow x\sigma(\underline{P})y.$$

In other words we aggregate the agents' preferences according to some social welfare function satisfying all Arrovian rationality axioms, and this aggregation yields a single preference relation which belongs to the government.

The social welfare function induces a collection of so-called *σ-decisive coalitions*. A coalition C is said to be σ-decisive if and only if for all $x, y \in \mathcal{C}$ and $\underline{P} = (P^n)_{n \in N} \in \mathcal{P}^N$ one has $x\sigma(\underline{P})y$, whenever $xP^n y$, for all $n \in D$. We denote the set of all σ-decisive coalitions by F_σ. Since our economy consists of an infinite number of agents and the social welfare function satisfies all Arrovian rationality axioms, F_σ is a non-principal ultrafilter (see Section 3.2).

4.2.3 Agent's happiness function and assumptions

Agents in the economy maximise their *happiness* which is a function that depends on the social and exogenous parameters, as well as their choices. We impose some assumptions which are crucial for our results.

At first we formulate the choice set of the agent as follows:

$$\mathcal{Y} = \left\{ (b_t, m_t, r_t, \pi_t)_t : c_t + m_t + b_t = (1 + r_{t-1})b_{t-1} + \frac{m_{t-1}}{(1 + \pi_t)} \right\}, \tag{4.11}$$

where $(c_t)_t \in X$, $(b_t, m_t)_t \in \mathcal{A}$ and $(r_t, \pi_t)_t \in \mathcal{C}$.

The interpretation of the choice set $\mathcal{Y}$ is that the agent chooses paths for the number of real bonds, real money holdings, real rate of interest and the inflation rate subject to his (or her) budget constraint (4.10).

Assumption 4.4. *The happiness functions of the agents are of the form $f : X \times \mathcal{Y} \to \mathbb{R}$, where X is as in Section 4.2.1.1 and $\mathcal{Y}$ is as* (4.11)*.*

We impose the following additional assumptions on the choice set and the happiness functions:

Assumption 4.5. *We assume that $\mathcal{Y}$ is a compact non-empty convex subset of $\mathcal{A} \times \mathcal{C}$, where $\mathcal{A}$ and $\mathcal{C}$ are as in Section 4.2.1. Furthermore every happiness function $f : X \times \mathcal{Y} \to \mathbb{R}$ is continuous and $f(x, \cdot)$ is strictly concave for all $x \in X$.*

The maximisation occurs subject to a budget constraint:

Assumption 4.6. *Let $\mathcal{Y}$ be as in Assumption 4.5. We assume*

$$\mathcal{M} \subseteq \left\{ \max_{(b_t,m_t)_t:(b_t,m_t,\cdot)_t \in \mathcal{Y}} f\left((c_t)_t, (b_t, m_t)_t, \cdot\right) : \begin{array}{c} f \text{ satisfies Assumptions 4.4 and 4.5,} \\ (c_t)_t \in X \end{array} \right\},$$

where $\mathcal{M}$, as we have discussed in Section 4.2.2, is some class of functions from $\mathcal{C}$ to $\mathbb{R}$ such that every agent's utility function belongs to $\mathcal{M}$.

It is obvious that for every *utility profile* $\underline{u} = (u^n)_{n \in N}$ in $\mathcal{M}^N$, there are N-sequences $(f^n)_{n \in N}$ and $((c^n_t)_t)_{n \in N} \in X^N$ such that

$$u^n = \max_{(b^n_t,m^n_t)_t:(b^n_t,m^n_t,\cdot)_t \in \mathcal{Y}} f^n\left((c^n_t)_t, (b^n_t, m^n_t)_t, \cdot\right),$$

for every $n \in N$.

4.2.4 A socially optimal path for the economy

The agents and the government enter a period t holding asset stocks dated $t-1$. As we have mentioned before, the government finances its expenditures by borrowing from the agents and printing money; at the same time, the agents consume and decide what level of real bonds and real money holdings (dated t) to carry over to the start of period $t+1$. Thus the government is going to maximise a *representative utility function* subject to its choices, and most importantly, subject to the agents' demand for money and bonds. Therefore we can define a *socially optimal path for the economy*, but before the definition, we introduce the following abbreviations:

Notation 4.7. *Denote $\vec{c} = (c_t)_t$, $\vec{r} = (r_t)_t$, $\vec{\pi} = (\pi_t)_t$, $\vec{b} = (b_t)_t$ and $\vec{m} = (m_t)_t$, and denote the path profiles for the bond and money by $\underline{\vec{b}}$ and $\underline{\vec{m}}$ respectively.*

Now we have the following definition:

Definition 4.8. A *σ-socially optimal path for the economy* is a quadruple $(\underline{\vec{b}}, \underline{\vec{m}}, \vec{r}, \vec{\pi}) \in \mathcal{A}^N \times \mathcal{C}$ such that for all $n \in N$, $(\vec{b^n}, \vec{m^n}, \vec{r}, \vec{\pi}) \in \mathcal{Y}$ and

1. Government maximisation: $(\vec{r}, \vec{\pi})$ maximises the σ-representative utility function subject to $(\vec{r}, \vec{\pi}) \in \mathcal{C}$.

2. Agent maximisation: For all $n \in N$, $(\vec{b^n}, \vec{m^n})$ maximises the happiness function f^n subject to the budget constraint (4.10).

Theorem 4.9. *Suppose Assumptions 4.4, 4.5 and 4.6 hold and the social welfare function σ satisfies all Arrovian rationality axioms. Then there exists a σ-socially optimal path for the economy.*

Proof. The assumptions in this theorem satisfy the conditions of Corollary 4.3 in Chapter 3. Therefore there exists a σ-representative utility function $\bar{u}$ for the utility profile $\underline{u}$. But according to the definition of the σ-representative utility function (see Definition 3.22),

$$\exists(\vec{r},\vec{\pi}) \in \arg\max_{(\vec{r}.\vec{\pi})\in\mathcal{C}} \bar{u}.$$

Assumption 4.6 holds, then for all $n \in N$ and every utility profile $\underline{u}(\vec{r},\vec{\pi}) = (u^n(\vec{r},\vec{\pi}))_{n\in N}$ in $\mathcal{M}^N$, there exist N-sequences $(f^n)_{n\in N}$ and $\underline{\vec{c}} \in X^N$ such that

$$u^n(\vec{r},\vec{\pi}) = \max_{(\vec{b^n},\vec{m^n}):(\vec{b^n},\vec{m^n},\vec{r},\vec{\pi})\in\mathcal{Y}} f^n(\vec{c^n},\vec{b^n},\vec{m^n},\vec{r},\vec{\pi}).$$

On the other hand f^n is continuous and $X \times \mathcal{Y}$ is compact, therefore $f^n(X \times \mathcal{Y})$ is a compact subset of $\mathbb{R}$ which contains its least upper bound. This proves the existence of a maximiser $(\vec{b^n},\vec{m^n})$ for every happiness function f^n subject to the budget constraint (4.10), hence, the proof is complete. □

Remark 4.10. (Socially Optimal Seigniorage). The government needs to generate revenue for financing its consumptions and debts, and one way for reaching this goal is printing money. The real government revenue from printing money is called *seigniorage* and is given by equation (4.9). The government budget constraint implies that in the case of any change in seigniorage, the government has to adjust the other parameters of the budget constraint in order to rebalance it. For instance, reducing the inflation rate to zero implies that the government must replace the seigniorage revenue by an increase in real borrowings or reduce the expenditures. But seigniorage as a function of inflation faces a Laffer curve (see Walsh [62]), which says that inflation raising after a certain point would cause seigniorage to reduce. The government has a strong incentive to be at that certain equilibrium point, because it is the point at which the government collects maximum amount of seigniorage. The socially optimal inflation path that we have found in Theorem 4.9 is the equilibrium point for the inflation.

4.2.5 Special cases

Every agent n maximises the happiness function f^n subject to the budget constraint (4.10). For solving the agent's maximisation problem, we impose the following assumption in addition to Assumptions 4.4, 4.5 and 4.6.

Assumption 4.11. *Each happiness function $f : X \times \mathcal{Y} \to \mathbb{R}$ is differentiable.*

We are confronted with the maximisation problem:

$$\max_{b_t^n, m_t^n} f^n(c_t^n, b_t^n, m_t^n, r_t, \pi_t) \tag{4.12}$$

subject to

$$x_t^n \equiv (1+r_{t-1})b_{t-1}^n + \frac{m_{t-1}^n}{(1+\pi_t)} - c_t^n = b_t^n + m_t^n, \tag{4.13}$$

and

$$x_{t+1}^n = (1+r_t)b_t^n + \frac{m_t^n}{(1+\pi_{t+1})} - c_{t+1}^n, \tag{4.14}$$

where (4.13) is the inter-temporal budget constraint at time t and (4.14) is the inter-temporal budget constraint at time $t+1$.

To maximise the system of equations (4.12)–(4.14), we can apply the method of Lagrangian multiplier to solve the model:

$$\begin{aligned} L = f^n(c_t^n, b_t^n, m_t^n, r_t, \pi_t) &+ \lambda_t(x_t^n - b_t^n - m_t^n) \\ &+ \lambda_{t+1}\left(x_{t+1}^n - (1+r_t)b_t^n - \frac{m_t^n}{(1+\pi_{t+1})} + c_{t+1}^n\right), \end{aligned} \tag{4.15}$$

where λ_t and λ_{t+1} are the Lagrangian multipliers. The first–order necessary conditions for this problem are:

$$\frac{\partial L}{\partial b_t} = 0 \Rightarrow f_{b_t}^n(c_t^n, b_t^n, m_t^n, r_t, \pi_t) - \lambda_t - \lambda_{t+1}(1+r_t) = 0 \tag{4.16}$$

$$\frac{\partial L}{\partial m_t} = 0 \Rightarrow f_{m_t}^n(c_t^n, b_t^n, m_t^n, r_t, \pi_t) - \lambda_t - \lambda_{t+1}\left(\frac{1}{1+\pi_{t+1}}\right) = 0 \tag{4.17}$$

$$\frac{\partial L}{\partial x_t^n} = 0 \Rightarrow \lambda_t = 0 \tag{4.18}$$

$$\frac{\partial L}{\partial \lambda_t} = 0 \Rightarrow x_t^n = b_t^n + m_t^n \tag{4.19}$$

$$\frac{\partial L}{\partial \lambda_{t+1}} = 0 \Rightarrow x_{t+1}^n = (1+r_t)b_t^n + \frac{m_t^n}{(1+\pi_{t+1})} - c_{t+1}^n \tag{4.20}$$

From equations (4.16)–(4.18), we have:

$$\begin{aligned} \frac{f_{m_t}^n(c_t^n, b_t^n, m_t^n, r_t, \pi_t)}{f_{b_t}^n(c_t^n, b_t^n, m_t^n, r_t, \pi_t)} &= \frac{1}{(1+\pi_{t+1})(1+r_t)} \\ &= \frac{1}{1+i_t}, \end{aligned} \tag{4.21}$$

where $1+i_t = (1+r_t)(1+\pi_{t+1})$ (by the definition of r_t).

The equation (4.21) states the relation between marginal benefit of additional money holdings and the marginal benefit of additional bond holdings of every agent n at time t. In other words, the marginal rate of substitution between money and bond for each agent n equals to $\frac{1}{1+i_t}$.

Now we are interested to set the real interest and inflation rates in a steady state, to maximise the utility of the representative agent, subject to the government revenue requirement. The utility of the representative agent depends on the real interest and inflation rates, $\bar{u}(r, \pi)$. The government's budget constraint in the steady state is

$$g = \left(\frac{\pi}{1+\pi}\right) m - rb. \tag{4.22}$$

We are concerned with the maximisation problem

$$\max_{r,\pi} \bar{u}(r, \pi), \tag{4.23}$$

subject to (4.22). We can apply again the method of Lagrangian multipliers to solve this problem:

$$L = \bar{u}(r, \pi) + \lambda \left(g - \left(\frac{\pi}{1+\pi}\right) m + rb\right),$$

where λ is the Lagrangian multiplier. The first–order necessary conditions for this problem are:

$$\frac{\partial L}{\partial r} = 0 \Rightarrow \bar{u}_r + \lambda b = 0 \tag{4.24}$$

$$\frac{\partial L}{\partial \pi} = 0 \Rightarrow \bar{u}_\pi - \lambda m \left(\frac{1}{(1+\pi)^2}\right) = 0 \tag{4.25}$$

$$\frac{\partial L}{\partial \lambda} = 0 \Rightarrow g - \left(\frac{\pi}{1+\pi}\right) m + rb = 0 \tag{4.26}$$

From equations (4.24)–(4.26), we have:

$$\frac{\bar{u}_r}{\bar{u}_\pi} = -\frac{b(1+\pi)^2}{m}, \tag{4.27}$$

where $-\frac{b(1+\pi)^2}{m}$ is the marginal rate of substitution between real interest and inflation rates.

Chapter 5

Generalisation to the case of weak compactness

This chapter builds on a requirement for existence of a representative agent model for weaker conditions. In Chapter 3, we have constructed a representative utility function for infinite-dimensional social decision problems and have assumed that the set of social alternatives is compact. Therefore, we establish weaker conditions for our previous results by allowing the social alternatives to be a weakly compact subset of a given reflexive separable Banach space. At first, we introduce a new and simple nonstandard approach to weak topological spaces and then prove the existence of a representative utility function with respect to this new constructed weak topology.

5.1 Nonstandard characterisation of weak compactness

In this section we try to present three kinds of notions and results: (a) the nonstandard universe for pseudo-normed linear spaces and the properties of the nonstandard hull with respect to this new nonstandard universe; (b) the weak topology on a given normed linear space by some pseudo-norms on this space and the properties of nonstandard hull with respect to this new weak topology; (c) some important results in weak compactness.

5.1.1 Pseudo-norms and nonstandard hulls

In this section we follow the notations from Albeverio [2] and Ng [49]. Let X be a vector space over $\mathbb{R}$. According to Sharma and Vasishtha in [57] and Deutsch in [17], a *pseudo-norm* $\|\cdot\|$ on X is a function $\|\cdot\| : X \to \mathbb{R}$ such that for all $x, y \in X$ and $c \in \mathbb{R}$ we have:

1. $\|cx\| = |c|\|x\|$, (Absolute homogeneity)
2. $\|x + y\| \leq \|x\| + \|y\|$. (Triangle inequality)

Proposition 5.1. *1. (Positivity). If absolute homogeneity and triangle inequality hold, then for all $x \in X$, $\|x\| \geq 0$.*

2. *A pseudo-norm that satisfies*

$$\forall x \in X \quad \|x\| = 0 \Rightarrow x = 0, \qquad \text{(Separating points)}$$

is a norm.

Proof. 1. By the absolute homogeneity we have $\|0\| = 0$ and $\|-x\| = \|x\|$ for all $x \in X$, so by the triangle inequality,

$$0 = \|0\| = \|x - x\| = \|x + (-x)\| \leq \|x\| + \|-x\| = 2\|x\|,$$

hence $\|x\| \geq 0$.

2. $\|\cdot\| : X \to \mathbb{R}$ is a norm, if it satisfies positivity, absolute homogeneity, triangle inequality and separating points conditions. Since $\|\cdot\|$ is a pseudo-norm, by the first part of this proposition and assumption of the second part, $\|\cdot\|$ is a norm.

□

Let A be a subset of X, then for $x \in X$ and $c \in \mathbb{R}$,

$$x + A := \{x + y \,:\, y \in A\}$$

and

$$cA := \{cx \,:\, x \in A\}.$$

Given X and subspace $Y \subset X$, we define

$$X/Y := \{x + Y \,:\, x \in X\}.$$

For $x, y \in X$ and $c \in \mathbb{R}$, the operation

$$(x + Y) + c(y + Y) := (x + cy) + Y,$$

is well-defined and X/Y form a vector space under this operation. We call this space as *quotient space.*

The ultrapower construction can easily be adapted to construct an embedding

$$^* : V\left(X \cup \mathbb{R}\right) \longrightarrow V\left({}^*X \cup {}^*\mathbb{R}\right),$$

where ${}^*X$ is a vector space over ${}^*\mathbb{R}$, satisfying Extension and Transfer Principles. Henceforth, we work in the resulting nonstandard universe. In the nonstandard universe, ${}^*\|\cdot\| : {}^*X \longrightarrow {}^*\mathbb{R}$ is well-defined and by the Transfer Principle, properties 1 and 2 hold for all $x, y \in {}^*X$ and $c \in {}^*\mathbb{R}$.

Definition 5.2. An element $x \in {}^*X$ is called *finite* if ${}^*\|x\|$ is a finite hyperreal.

We let $\mathrm{Fin}({}^*X)$ denote the finite elements of ${}^*X$ and we have

$$\mathrm{Fin}({}^*X) := \{x \in {}^*X \,:\, {}^*\|x\| \in \mathrm{Fin}({}^*\mathbb{R})\}.$$

Definition 5.3. We call the element $x \in {}^*X$ a *infinitesimal* if ${}^*\|x\| \approx 0$.

For any elements $x, y \in {}^*X$, we shall write $x \approx y$ to mean that the difference $x - y$ is infinitesimal.

Proposition 5.4. *If $x, y \in {}^*X$ and $x \approx y$, then ${}^*\|x\| \approx {}^*\|y\|$.*

Proof. Suppose that ${}^*\|x\| \leq {}^*\|y\|$. Then

$${}^*\|y\| = {}^*\|x + (y - x)\| \leq {}^*\|x\| + {}^*\|y - x\|.$$

So,

$$0 \leq {}^*\|y\| - {}^*\|x\| \leq {}^*\|y - x\| \approx 0,$$

therefore,

$${}^*\|y\| - {}^*\|x\| \approx 0,$$

that is ${}^*\|x\| \approx {}^*\|y\|$. □

Definition 5.5. An element $x \in {}^*X$ is called *nearstandard* if ${}^*\|x - y\| \approx 0$ for some $y \in X$.

Now we give the definition of the monad of an element in ${}^*X$:

Definition 5.6. For any $x \in {}^*X$, the *monad* of x is given by

$$\begin{aligned} \mu(x) &= \{y \in {}^*X \,:\, {}^*\|y - x\| \approx 0\} \\ &= \{y \in {}^*X \,:\, y \approx x\}. \end{aligned} \tag{5.1}$$

One can write

$$\mu(0) = \{y \in {}^*X \,:\, {}^*\|y\| \approx 0\}.$$

Remark 5.7. Both $\mathrm{Fin}({}^*X)$ and $\mu(0)$ are vector spaces over $\mathbb{R}$.

The notion of *nonstandard hull* was introduced for the first time by Luxemburg in [40]. He considered the case when X is a linear normed space. Later, Henson and Moore in [29] introduced this notion for Banach spaces. But, we study the nonstandard hull for the pseudo-norm linear spaces.

Definition 5.8. The quotient space

$$\hat{X} = \text{Fin}({}^*X)/\mu(0), \tag{5.2}$$

is called the *nonstandard hull of* X.

The elements of $\hat{X}$ are $x + \mu(0)$, where $x \in \text{Fin}({}^*X)$ and we have

$$\begin{aligned} x + \mu(0) &= \{x + y \,:\, y \in \mu(0)\} \\ &= \mu(x). \end{aligned} \tag{5.3}$$

Therefore, the nonstandard hull of X can be defined as follows:

$$\hat{X} = \{\mu(x) \,:\, x \in \text{Fin}({}^*X)\}. \tag{5.4}$$

Henceforth, we work with saturated embeddings satisfying Extension, Transfer and Internal Definition Principles, and suppose $(W, \|\cdot\|)$ is an internal pseudo-normed linear space. In this case we assume that the pseudo-norm is a map from W to ${}^*\mathbb{R}$ and the notions of finite, infinitesimal and monad still make sense. Also $\text{Fin}(W)$ and $\hat{W} = \text{Fin}(W)/\mu(0)$ are well-defined.

Proposition 5.9. *Suppose* $(W, \|\cdot\|)$ *is an internal pseudo-normed linear space. Then* $\hat{W}$ *(the nonstandard hull of* W*) is complete, i.e. is a Banach space.*

Proof. At first we should show that $\hat{W}$ is a normed linear space. Recall that

$$\hat{W} = \{\mu(w) \,:\, w \in \text{Fin}(W)\}.$$

For convenience $\mu(w)$ is denoted by $\hat{w}$ for every $w \in \text{Fin}(W)$, and $\mu(0)$ is denoted by $\hat{0}$. We define $\|\hat{w}\| := {}^\circ\|w\|$, for each $w \in \text{Fin}(W)$. It is easy to see that $\|\hat{w}\|$ is well-defined. Fix arbitrary $w, w' \in \text{Fin}(W)$ such that $\hat{w} = \hat{w}'$. Then $\mu(w) = \mu(w')$, therefore $w \approx w'$ and by Proposition 5.4, $\|w\| \approx \|w'\|$. Since $\|w\|, \|w'\| \in \text{Fin}({}^*\mathbb{R})$ and every finite element in ${}^*\mathbb{R}$ has a unique standard part (see Proposition 2.56), thus the standard part of $\|w\|$ and $\|w'\|$ are equal. Therefore $\|\hat{w}\|$ is well-defined.

The needed properties of a normed linear space for all $w, v \in \text{Fin}(W)$ and $c \in \text{Fin}({}^*\mathbb{R})$ are:

1. $\|\hat{w}\| \approx \|w\| \geq 0.$

2. $\|c\hat{w}\| = \|\widehat{cw}\| \approx \|cw\| = |c|\|w\| \approx |c|\|\hat{w}\|$. Therefore,

$$\|c\hat{w}\| = |c|\|\hat{w}\|.$$

3. $\|\hat{w}+\hat{v}\| = \|\widehat{w+v}\| \approx \|w+v\| \leq \|w\| + \|v\| \approx \|\hat{w}\| + \|\hat{v}\|$. Therefore,

$$\|\hat{w}+\hat{v}\| \leq \|\hat{w}\| + \|\hat{v}\|.$$

4. $0 = \|\hat{w}\| \approx \|w\| \Rightarrow \|w\| \approx 0$, which means that $w \in \mathrm{Fin}(W)$ is infinitesimal, therefore

$$\hat{w} = \mu(0) = \hat{0}.$$

Now we prove the completeness of $\hat{W}$. Let $(\hat{w}_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\hat{W}$, where $\hat{w}_n = \mu(w_n)$ for every $n \in \mathbb{N}$. We put

$$A_n = \{f \mid f : {}^*\mathbb{N} \to W\,, \quad f \text{ is } {}^*\text{Cauchy}\,, \quad \forall i \leq n \quad (f_i = w_i)\},$$

for every $n \in \mathbb{N}$. By the Internal Definition Principle (Proposition 2.76), the A_n's are non-empty and internal. Moreover the sequence $(A_n)_{n\in\mathbb{N}}$ is $\subseteq$-decreasing and hence by the Countable Saturation Principle (Proposition 2.79), the intersection

$$A = \bigcap_{n\in\mathbb{N}} A_n$$

is non-empty. This means that we can extend the sequence $(w_n)_{n\in\mathbb{N}}$ in W to an internal *Cauchy sequence $(w_n)_{n\in{}^*\mathbb{N}}$ in W.

For any $k \in \mathbb{N}^+$, there exists an $N \in \mathbb{N}$ such that for all $m, n \in \mathbb{N}$ with $m, n > N$,

$$\|w_n - w_m\| \approx \|\hat{w}_n - \hat{w}_m\| < \frac{1}{k},$$

hence $\|w_n - w_m\| < \frac{1}{k}$. The set

$$\{M \in {}^*\mathbb{N} : \forall m, n \in [N, M] \quad \|w_n - w_m\| < \frac{1}{k}\}$$

is internal and $\{N, N+1, \dots\} \subseteq \{M \in {}^*\mathbb{N} : \forall m, n \in [N, M] \quad \|w_n - w_m\| < \frac{1}{k}\}$. By using the Overspill Principle (Proposition 2.77), there exists some $M_k \in {}^*\mathbb{N}\backslash\mathbb{N}$ such that for all $N < n < m < M_k$ we have

$$\|w_n - w_m\| \leq \frac{1}{k},$$

therefore for all $m, n \in (N, M_k)$, $\|\hat{w}_n - \hat{w}_m\| \leq \frac{1}{k}$. Hence, $w_m \in \mathrm{Fin}(W)$ for all $m < M_k$. Now let $\alpha \in {}^*\mathbb{N}\backslash\mathbb{N}$ and $\alpha < M_k$. Then $w_\alpha \in \mathrm{Fin}(W)$ and

$$\lim_{n\to\infty} \|\hat{w}_n - \hat{w}_\alpha\| = 0,$$

hence $\hat{w}_n \to \hat{w}_\alpha \in \hat{W}$ and this completes the proof of the proposition. □

5.1.2 Nonstandard hull with respect to the weak topology

Let X be a set and $(f_i)_{i\in I}$, where I is an index set, be a collection of maps such that each f_i maps X into $\mathbb{R}$. We wish to define a topology on X that makes all the f_i's continuous, and we want to do this in the cheapest way, that is, there should be no more open sets in X than required for this purpose.

Obviously, all the $f_i^{-1}(O)$, where O is an open set in $\mathbb{R}$ should be open in X. Then finite intersections of those should also be open. And then any union of finite intersections should be open. By this process, we have created as few open sets as required.

Now let $\mathscr{O} \subset \mathcal{P}(X)$ be a collection of subsets of X such that

- $\emptyset$ and X are in $\mathscr{O}$;
- $\mathscr{O}$ is closed under finite intersections.

Then $\mathscr{T} = \{\bigcup_{O\in\mathcal{O}} O \,:\, \mathcal{O} \subset \mathscr{O}\}$ is a topology on X. Therefore the collection of all unions of finite intersection of sets of the form $f_i^{-1}(O)$ where $i \in I$ and O is an open set in $\mathbb{R}$ is a topology. It is called the *weak topology* on X generated by the $(f_i)_{i\in I}$'s and we denote it by $\mathscr{T}^w$. Also the functions $(f_i)_{i\in I}$ are continuous for this topology.

We want to generate a weak topology on X by some pseudo-norms on X. For this purpose, we suppose that X is a normed linear space over $\mathbb{R}$ with norm $\|\cdot\|_X$, and let X' be the dual of X (the set of all bounded linear functionals on X). For each $x \in X$ and $\phi \in X'$, the mapping $\|\cdot\|_\phi : X \to \mathbb{R}$ given by

$$\|x\|_\phi := \frac{1}{\|\phi\|}|\phi(x)|, \tag{5.5}$$

is a pseudo-norm on X, since for every $x, y \in X$ and $c \in \mathbb{R}$:

1. $\|cx\|_\phi = \frac{1}{\|\phi\|}|\phi(cx)| = \frac{1}{\|\phi\|}|c\phi(x)| = |c|\|x\|_\phi$,
2. $\|x+y\|_\phi = \frac{1}{\|\phi\|}|\phi(x+y)| = \frac{1}{\|\phi\|}|\phi(x)+\phi(y)| \le \frac{1}{\|\phi\|}\left(|\phi(x)| + |\phi(y)|\right) = \|x\|_\phi + \|y\|_\phi$.

Henceforth, we shall understand the *weak topology* on X to be the weak topology generated by $(\|\cdot\|_\phi)_{\phi\in X'}$.

Definition 5.10. A sequence $(x_n)_{n\in\mathbb{N}}$ in X *converges weakly* to $x \in X$, if for all $\phi \in X'$ we have:

$$\lim_{n\to\infty} \|x_n - x\|_\phi = 0.$$

In this case we write $x_n \xrightarrow{w} x$.

As in the previous section, we construct an embedding

$$^* : V\left(X \cup \mathbb{R}\right) \longrightarrow V\left({}^*X \cup {}^*\mathbb{R}\right),$$

where ${}^*X$ is a normed linear space over ${}^*\mathbb{R}$ satisfying the Extension, Transfer and Internal Definition Principles. In the nonstandard universe, ${}^*\|\cdot\|_\phi : {}^*X \to {}^*\mathbb{R}$ is well-defined and by the Transfer Principle conditions 1 and 2 hold for any $x, y \in {}^*X$ and $c \in {}^*\mathbb{R}$. We define the infinitely close relation with respect to the weak topology as follows:

Definition 5.11. For any element $x, y \in {}^*X$, we write $x \approx_w y$ to mean that

$$\forall \phi \in {}^*X' \qquad \frac{1}{\|\phi\|}|\phi(x) - \phi(y)| \approx 0.$$

Definition 5.12. We call the element $x \in {}^*X$ a *infinitesimal with respect to the weak topology* if $x \approx_w 0$.

We can define the monad of an element in the weak topology as follows:

Definition 5.13. For any $x \in {}^*X$, the *monad of x w.r.t. the weak topology* is given by

$$\mu_w(x) = \{y \in {}^*X \,:\, x \approx_w y\}.$$

Remark 5.14. For any $x \in {}^*X$,

$$\mu_w(x) = \bigcap_{\delta \in \mathbb{R}_{>0}} \{y \in {}^*X \,:\, \forall \phi \in {}^*X',\, \|y - x\|_\phi < \delta\}.$$

Proof. Take an arbitrary $y \in \mu_w(x)$. Since $y \in \mu_w(x)$, $x \approx_w y$. That is,

$$\forall \phi \in {}^*X' \qquad \frac{1}{\|\phi\|}|\phi(x) - \phi(y)| = \frac{1}{\|\phi\|}|\phi(x - y)| = \|x - y\|_\phi \approx 0.$$

This means that for all $n \in \mathbb{N}\backslash\{0\}$, $\|x - y\|_\phi < \frac{1}{n}$. Hence by taking $\delta = \frac{1}{n}$,

$$\mu_w(x) \subseteq \bigcap_{n \in \mathbb{N}\backslash\{0\}} \{y \in {}^*X \,:\, \forall \phi \in {}^*X',\, \|y - x\|_\phi < \frac{1}{n}\}.$$

Now suppose that $y \in \bigcap_{\delta \in \mathbb{R}_{>0}} \{y \in {}^*X \,:\, \forall \phi \in {}^*X',\, \|y - x\|_\phi < \delta\}$. Hence, for every $\delta \in \mathbb{R}_{>0}$, we have

$$\{y \in {}^*X \,:\, \forall \phi \in {}^*X',\, \|y - x\|_\phi < \delta\},$$

for any $x \in {}^*X$. That is, for all $\phi \in {}^*X'$,

$$\delta > \|y - x\|_\phi = \frac{1}{\|\phi\|}|\phi(x - y)| = \frac{1}{\|\phi\|}|\phi(x) - \phi(y)|.$$

Since this is true for every δ, we can take $\delta = \frac{1}{n}$ for every $n \in \mathbb{N}\backslash\{0\}$, which means that $y \in \mu_w(x)$. $\square$

Let $S_{X'}$ be the unit sphere in X', defined by

$$S_{X'} := \{\phi \in X' \,:\, \|\phi\|_{X'} = 1\},$$

where $\|\phi\|_{X'} = \sup\{|\phi(x)| \,:\, x \in X\,,\, \|x\|_X = 1\}$. Then we have the following definition:

Definition 5.15. An element $x \in {}^*X$ is called *finite w.r.t. the weak topology* if $\sup_{\phi\in {}^*S_{X'}}|\phi(x)|$ is a finite hyperreal.

We let $\mathrm{Fin}_w({}^*X)$ denote the finite elements of ${}^*X$ in the weak topology, and we have

$$\mathrm{Fin}_w({}^*X) := \{x \in {}^*X \,:\, \sup_{\phi\in {}^*S_{X'}} {}^{\circ}|\phi(x)| < \infty\}.$$

Remark 5.16. As in Section 5.1.1, for every $x \in {}^*X$ we have $\mu_w(x) = x + \mu_w(0)$. Also for every $x \in \mathrm{Fin}_w({}^*X)$ and $y \in {}^*X$, if $x \approx_w y$ then $y \in \mathrm{Fin}_w({}^*X)$.

Definition 5.17. The quotient space

$$\begin{aligned}\hat{X}_w &= \mathrm{Fin}_w({}^*X)/\mu_w(0)\\ &= \{\mu_w(x) \,:\, x \in \mathrm{Fin}_w({}^*X)\}.\end{aligned} \tag{5.6}$$

is called the *nonstandard hull of X w.r.t. the weak topology.*

Consider an internal linear space B and let $\|\cdot\|_\phi$ be a pseudo-norm on B for every $\phi \in B'$, defined as (5.5). In this case the pseudo-norm is a map from B to ${}^*\mathbb{R}$ and the notions of finite, infinitesimal and monad w.r.t. the weak topology generated by $(\|\cdot\|_\phi)_{\phi\in B'}$ still make sense. Also $\mathrm{Fin}_w(B)$ and $\hat{B}_w = \mathrm{Fin}(B)/\mu(0)$ are well-defined

Proposition 5.18. *Suppose $(B, \|\cdot\|_\phi)$ is an internal pseudo-normed linear space for every $\phi \in S_{B'}$. Then $\hat{B}_w$ is complete (forms a Banach space).*

Proof. The proof is similar to the proof of Proposition 5.9. For convenience $\mu_w(b)$ is denoted by $\hat{b}$ for every $b \in \mathrm{Fin}_w(B)$, and $\mu_w(0)$ is denoted by $\hat{0}$. Let $\|\cdot\|_w : \hat{B}_w \to \mathbb{R}$ be a functional on $\hat{B}_w$ which is defined for every $\hat{b} \in \hat{B}_w$ by

$$\|\hat{b}\|_w := \sup_{\phi\in S_{B'}} {}^{\circ}|\phi(b)|.$$

At first we should show that $\|\cdot\|_w$ is a norm on $\hat{B}_w$. The needed properties of a normed linear space for all $\hat{a}, \hat{b} \in \hat{B}_w$ and $c \in \mathbb{R}$ are:

1. (Positivity). $\|\hat{b}\|_w = \sup_{\phi\in S_{B'}} {}^{\circ}|\phi(b)| \geq 0.$

2. (Absolute homogeneity).

$$\begin{aligned}\|c\hat{b}\|_w = \sup_{\phi\in S_{B'}} {}^\circ|\phi(cb)| &= \sup_{\phi\in S_{B'}} {}^\circ|c\phi(b)| \\ &= {}^\circ|c| \sup_{\phi\in S_{B'}} {}^\circ|\phi(b)| \\ &= |c| \sup_{\phi\in S_{B'}} {}^\circ|\phi(b)| \\ &= |c|\|\hat{b}\|_w.\end{aligned}$$

Therefore $\|c\hat{b}\|_w = |c|\|\hat{b}\|_w$.

3. (Triangle inequality). For all $\phi \in S_{B'}$, one can write

$$\begin{aligned}{}^\circ|\phi(a+b)| = {}^\circ|\phi(a)+\phi(b)| &\leq {}^\circ|\phi(a)| + {}^\circ|\phi(b)| \\ &\leq \sup_{\phi\in S_{B'}} {}^\circ|\phi(a)| + \sup_{\phi\in S_{B'}} {}^\circ|\phi(b)| \\ &= \|\hat{a}\|_w + \|\hat{b}\|_w,\end{aligned}$$

hence

$${}^\circ|\phi(a+b)| \leq \|\hat{a}\|_w + \|\hat{b}\|_w.$$

Since the right hand side of the inequality above is independent of ϕ, taking supremum of both sides of the inequality over ϕ yields:

$$\sup_{\phi\in S_{B'}} {}^\circ|\phi(a+b)| \leq \|\hat{a}\|_w + \|\hat{b}\|_w,$$

and therefore

$$\|\hat{a}+\hat{b}\|_w \leq \|\hat{a}\|_w + \|\hat{b}\|_w.$$

4. (Separating points). If for every $a \in \mathrm{Fin}_w(B)$, we have $0 = \|\hat{a}\|_w = \sup_{\phi\in S_{B'}} {}^\circ|\phi(a)|$, then

$$\forall\phi \in S_{B'} \qquad \phi(a) = 0,$$

which shows that $a \in \mathrm{Fin}_w(B)$ is an infinitesimal w.r.t. the weak topology. Hence $\hat{a} = \mu_w(0) = \hat{0}$.

Now we prove the completeness of $\hat{B}_w$. Let $(\hat{b}_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $\hat{B}_w$. For each $n \in \mathbb{N}$ and $\phi \in S_{B'}$, we define the set

$$A_{n,\phi} = \{a \,|\, a : {}^*\mathbb{N} \to B\,, \quad a \text{ is } {}^*\text{Cauchy}\,, \quad \forall i \leq n \quad (\phi(a_i) = \phi(b_i))\}\,.$$

By the Internal Definition Principle (Proposition 2.76), the $A_{n,\phi}$'s are non-empty and internal. Moreover the sequence $(A_{n,\phi})_{n\in\mathbb{N}}$ is $\subseteq$-decreasing and hence by the Countable

Saturation Principle (Proposition 2.79), the intersection

$$A = \bigcap_{n\in\mathbb{N},\,\phi\in S_{B'}} A_{n,\phi}$$

is non-empty. This means that we can extend the sequence $(b_n)_{n\in\mathbb{N}}$ in B to an internal *Cauchy sequence $(b_n)_{n\in{}^*\mathbb{N}}$ in B. For any $k \in \mathbb{N}^+$, there exists an $N \in \mathbb{N}$ such that for all $m, n \in \mathbb{N}$ with $m, n > N$,

$$\sup_{\phi\in S_{B'}} {}^\circ|\phi(b_n - b_m)| = \|\hat{b}_n - \hat{b}_m\|_w < \frac{1}{k},$$

hence $\sup_{\phi\in S_{B'}} {}^\circ|\phi(b_n - b_m)| < \frac{1}{k}$. The set

$$\{M \in {}^*\mathbb{N} : \forall m, n \in [N, M] \quad \sup_{\phi\in S_{B'}} {}^\circ|\phi(b_n - b_m)| < \frac{1}{k}\}$$

is internal and

$$\{N, N+1, \ldots\} \subseteq \{M \in {}^*\mathbb{N} : \forall m, n \in [N, M] \quad \sup_{\phi\in S_{B'}} {}^\circ|\phi(b_n - b_m)| < \frac{1}{k}\}.$$

By using the Overspill Principle (Proposition 2.77), there exists some $M_k \in {}^*\mathbb{N}\backslash\mathbb{N}$ such that for all $N < n < m < M_k$ we have

$$\sup_{\phi\in S_{B'}} {}^\circ|\phi(b_n - b_m)| \leq \frac{1}{k}.$$

Therefore for all $m, n \in (N, M_k)$, $\|\hat{b}_n - \hat{b}_m\|_w \leq \frac{1}{k}$. Hence, $b_m \in \mathrm{Fin}_w(B)$ for all $m < M_k$. Now let $\alpha \in {}^*\mathbb{N}\backslash\mathbb{N}$ and $\alpha < M_k$. Then $b_\alpha \in \mathrm{Fin}_w(B)$ and

$$\lim_{n\to\infty} \|\hat{b}_n - \hat{b}_\alpha\| = 0,$$

hence $\hat{b}_n \to \hat{b}_\alpha \in \hat{B}_w$ and this completes the proof of the proposition. □

5.1.3 Weak compactness

Let X and ϕ be as in Section 5.1.2 and consider the weak topology $\mathscr{T}^w$ on X generated by the pseudo-norms $(\|\cdot\|_\phi)_{\phi\in X'}$. We call the elements of $\mathscr{T}^w$ *weakly open sets.*

Remark 5.19. Let $\mathscr{T}^w_x$ be the family of weakly open sets containing $x \in X$. Then

$$\mu_w(x) = \bigcap\{{}^*T : T \in \mathscr{T}^w_x\}.$$

Proof. Suppose $y \in \mu_w(x)$. Due to Remark 5.14, for all $\delta \in \mathbb{R}_{>0}$ and all $\phi \in {}^*X'$, we have $\|y - x\|_\phi < \delta$. Now suppose that $T \in \mathscr{T}^w_x$. Then there exists $\delta \in \mathbb{R}_{>0}$ such that the formula

$$\forall \phi \in X' \qquad \|z-x\|_\phi < \delta \Rightarrow z \in T,$$

holds in $V(X \cup \mathbb{R})$. By the Transfer Principle,

$$\forall \phi \in {}^*X' \qquad \|z-x\|_\phi < \delta \Rightarrow z \in {}^*T,$$

holds in $V({}^*X \cup {}^*\mathbb{R})$. Since this holds for each T satisfying $T \in \mathscr{T}_x^w$,

$$y \in \bigcap \{ {}^*T : T \in \mathscr{T}_x^w \}.$$

Hence $\mu_w(x) \subseteq \bigcap \{ {}^*T : T \in \mathscr{T}_x^w \}$.

Conversely suppose that $y \in \bigcap \{ {}^*T : T \in \mathscr{T}_x^w \}$. Choose $\delta \in \mathbb{R}_{>0}$ and let

$$T = \{ z \in X : \forall \phi \in X', \|z-x\|_\phi < \delta \}.$$

$T \in \mathscr{T}_x^w$, so $y \in {}^*T$. Therefore for all $\phi \in {}^*X'$, $\|y-x\|_\phi < \delta$. Hence $y \in \mu_w(x)$. □

Definition 5.20. We call a point $x \in {}^*X$ *nearstandard w.r.t. the weak topology* if $x \in \mu_w(y)$ for some $y \in X$. According to our previous notations, an element $x \in {}^*X$ is nearstandard w.r.t. the weak topology if $x \approx_w y$ for some $y \in X$.

By using the idea from Royden and Fitzpatrick [54], we give the following definition:

Definition 5.21. A subset D of X is said to be *weakly dense in* X if every non-empty weakly open set contains a point of D. The space X is said to be *weakly separable* if there is a countable subset of X that is weakly dense in X.

Proposition 5.22. *Let X be weakly separable. Then there exists a countable collection $(T_i)_{i\in\mathbb{N}}$ of weakly open sets such that any weakly open set is the union of a sub-collection of $(T_i)_{i\in\mathbb{N}}$.*

Proof. Let D be a countable weakly dense subset of X. If D is finite, then $X = D$. Assume D is countably infinite. Let $(x_i)_{i\in\mathbb{N}}$ be an enumeration of D. Then $\left(B\left(x_n, \frac{1}{m}\right)\right)_{n,m\in\mathbb{N}}$ is a countable collection of weakly open sets of X. We claim that every weakly open set of X is the union of a sub-collection of $\left(B\left(x_n, \frac{1}{m}\right)\right)_{n,m\in\mathbb{N}}$. Indeed, let $T \in \mathscr{T}^w$ be a weakly open subset of X. Let $x \in T$. We must show that there are natural numbers n and m for which

$$x \in B\left(x_n, \frac{1}{m}\right) \subseteq T.$$

Since T is weakly open and $x \in T$, there is a natural number m for which $B\left(x, \frac{1}{m}\right) \subseteq T$. Since every weakly open set contains a point of D, we may choose a natural number n

for which $x_n \in D \cap B\left(x, \frac{1}{2m}\right)$. Thus

$$x \in B\left(x_n, \frac{1}{2m}\right) \subseteq B\left(x, \frac{1}{m}\right) \subseteq T$$

holds for this choice of n and m. □

Theorem 5.23. *Suppose X is weakly separable. X is weakly compact if and only if every element in ${}^*X$ is nearstandard w.r.t. the weak topology.*

Proof. Suppose X is weakly compact and there is some $y \in {}^*X$ which is not nearstandard w.r.t. the weak topology. Then for every $x \in X$, there exists T_x with $x \in T_x \in \mathscr{T}^w$ and $y \notin {}^*T_x$. Therefore $\{T_x : x \in X\}$ is a weakly open cover of X. Let $\{T_{x_0}, \ldots, T_{x_{n-1}}\}$ be a finite subcover of X for $n \in \mathbb{N}$. Hence

$$\bigcup_{i=0}^{n-1} {}^*T_{x_i} = {}^*\left(\bigcup_{i=0}^{n-1} T_{x_i}\right) = {}^*X,$$

so $y \notin {}^*X$, which is a contradiction.

Conversely, suppose that every $y \in {}^*X$ is nearstandard w.r.t. the weak topology. Let $\{T_\lambda : \lambda \in \Lambda\}$ be an weakly open cover of X. Since X is separable, due to Proposition 5.22, there exists a countable collection $(T_{\lambda_i})_{i\in\mathbb{N}}$ of weakly open sets such that any weakly open set T_λ, for $\lambda \in \Lambda$, is the union of a sub-collection of $(T_{\lambda_i})_{i\in\mathbb{N}}$. Define $C_{\lambda_i} := X \backslash T_{\lambda_i}$ for every $i \in \mathbb{N}$. If there is no finite subcover, then for every collection $\{\lambda_0, \ldots, \lambda_{n-1}\}$, $n \in \mathbb{N}$, we have

$$\bigcap_{i=0}^{n-1} C_{\lambda_i} \neq \emptyset.$$

Therefore

$$\bigcap_{i=0}^{n-1} {}^*C_{\lambda_i} = {}^*\left(\bigcap_{i=0}^{n-1} C_{\lambda_i}\right) \neq \emptyset.$$

By the Countable Saturation Principle 2.79,

$$C = \bigcap_{i\in\mathbb{N}} {}^*C_{\lambda_i} \neq \emptyset.$$

Given an arbitrary $x \in X$, there exist some $\lambda \in \Lambda$ and $N \in \mathbb{N}$ such that $x \in T_\lambda \subseteq T_{\lambda_N}$. Hence $x \in T_{\lambda_N}$. Let $y \in C$. Then $y \in C \subset {}^*C_{\lambda_N}$, therefore $y \notin {}^*T_{\lambda_N}$ and so $y \notin \mu_w(x)$. Because x is an arbitrary element of X, y is not nearstandard w.r.t. the weak topology and this is a contradiction. Thus, $\{T_\lambda : \lambda \in \Lambda\}$ has a finite subcover, and X is weakly compact. □

Theorem 5.24. *Suppose the embedding * is κ-saturated. X is weakly compact if and only if every element in ${}^*X$ is nearstandard w.r.t. the weak topology.*

Proof. If X is weakly compact, the proof of every element in ${}^*X$ is nearstandard w.r.t. the weak topology, is the same as the first part of the proof in Theorem 5.23.

Conversely, suppose that every $y \in {}^*X$ is nearstandard w.r.t. the weak topology. Let $\{T_\lambda : \lambda \in \Lambda\}$ be an weakly open cover of X. Define $C_\lambda := X \backslash T_\lambda$ for every $\lambda \in \Lambda$. If there is no finite subcover, then for every collection $\{\lambda_0, \ldots, \lambda_{n-1}\}$ with $n \in \mathbb{N}$, we have

$$\bigcap_{i=0}^{n-1} C_{\lambda_i} \neq \emptyset.$$

Therefore

$$\bigcap_{i=0}^{n-1} {}^*C_{\lambda_i} = {}^*\left(\bigcap_{i=0}^{n-1} C_{\lambda_i}\right) \neq \emptyset.$$

Since $|\Lambda| \leq |\mathcal{P}(X)| < \kappa$, where $\kappa \geq |V_1(X \cup \mathbb{R})|$, by κ-Saturation Principle (Definition 2.81),

$$C = \bigcap_{\lambda \in \Lambda} {}^*C_\lambda \neq \emptyset.$$

Given an arbitrary $x \in X$, there exists some λ such that $x \in T_\lambda$. Let $y \in C$. Then $y \in C \subset {}^*C_\lambda$, therefore $y \notin {}^*T_\lambda$ and so $y \notin \mu_w(x)$. Because x is an arbitrary element of X, y is not nearstandard w.r.t. the weak topology and this is a contradiction. Thus, $\{T_\lambda : \lambda \in \Lambda\}$ has a finite subcover, therefore X is weakly compact. □

5.2 Existence of a representative utility function w.r.t. the weak topology

In this section we are going to provide weaker conditions for the existence of a representative utility function. For this purpose, we assume weaker conditions for the set of social alternatives and the parameter set of parametrised utility functions. We use here the same notations and notions introduced in Chapter 3.

5.2.1 The model

Let N and $\mathcal{C}$ be the set of *individuals* and *social alternatives* respectively. There exists a *weak order* relation P on $\mathcal{C}$ and we denote the set of all weak orders by $\mathcal{P}$. Elements of $\mathcal{P}^N$ are called *profiles*. Every individual $i \in N$ has preferences that can be represented by a utility function $u_i : \mathcal{C} \to \mathbb{R}$. We fix some class $\mathcal{M}$ of functions from $\mathcal{C}$ to $\mathbb{R}$ such that every individual's utility function belongs to $\mathcal{M}$. Elements of $\mathcal{M}^N$ will be called *utility profiles* and $\underline{u} = (u_i)_{i \in N} \in \mathcal{M}^N$.

Recall that for $x \in \mathcal{C}$ and $P \in \mathcal{P}$, x will be called *P-maximal* if and only if for all $y \in \mathcal{C} \backslash \{x\}$, we have xPy.

Notation 5.25. *We have the following notations:*

- *For all* $x, y \in \mathcal{C}$ *and* $\underline{P} = (P_i)_{i \in N} \in \mathcal{P}^N$, *we define*

$$C(x, y, \underline{P}) := \{i \in N \,:\, xP_iy\}.$$

- *We use* $P^u \in \mathcal{P}$ *to show that the utility function* u *induces the preference* P.

- *Given an* N*-sequence* $\underline{u} = (u_i)_{i \in N}$ *of functions from* $\mathcal{C}$ *to* $\mathbb{R}$, *we define*

$$\underline{P}^{\underline{u}} := (P^{u_i})_{i \in N} \in \mathcal{P}^N.$$

We say that the utility profile $\underline{u}$ *induces the preference profile* $\underline{P}^{\underline{u}}$.

We aggregate the individuals' preferences with a *social welfare function*, $\sigma : \mathcal{P}^N \to \mathcal{P}$, which satisfies all Arrovian rationality axioms (Axioms 3.10, 3.11 and 3.12). A coalition $C \subseteq N$ is said to be σ-*decisive* if and only if for all $x, y \in \mathcal{C}$ and $\underline{P} \in \mathcal{P}^N$ one has $x\sigma(\underline{P})y$ whenever xP_iy for all $i \in C$. The set of σ-decisive coalitions is denoted by F_σ. We have shown in Chapter 3, if σ satisfies Arrovian rationality axioms, then F_σ is a non-principal ultrafilter on N and this is only possible if N is infinite.

According to our results, we impose the following assumption:

Assumption 5.26. *We assume that* $\mathcal{C}$ *is a weakly compact non-empty convex subset of a given reflexive separable Banach space* W *(with norm* $\|\cdot\|_W$*).*

Definition 5.27. Let $h : W \to \mathbb{R}$ be a function. We call the following limit, if it exists,

$$\lim_{\lambda \to 0} \frac{h(x + \lambda y) - h(u)}{\lambda},$$

the *directional derivative of* h *at* x *in the direction* y and denote it by $h'(x; y)$. If there exists $h_x \in W'$ (the dual space of W) such that

$$\forall y \in W \qquad h'(x; y) = \langle y, h_x \rangle\,,$$

we say that h is *Gâteaux-differentiable at* x, and h_x is called the *Gâteaux-differential of* h *at* x and we denote it by $h'(x)$.

Remark 5.28. The uniqueness of the Gâteaux-differential follows directly and it is characterised by

$$\forall y \in W \qquad \lim_{\lambda \to 0} \frac{h(x + \lambda y) - h(x)}{\lambda} = \big\langle y, h'(x) \big\rangle\,.$$

Definition 5.29. Let A be a weakly compact non-empty subset of a given separable Banach space X (with norm $\|\cdot\|_X$). The function $h : A \to \mathbb{R}$ is *w-uniformly continuous* *in* $x \in A$ if and only if

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall y \in A \quad \big((\forall \phi \in X' \quad \|x - y\|_\phi < \delta) \Rightarrow |h(x) - h(y)| < \epsilon \big) .$$

We have the following parametrisation assumption:

Assumption 5.30. *Let Z be a weakly compact non-empty subset of a given separable Banach space X. Let $v : Z \times \mathcal{C} \longrightarrow \mathbb{R}$ be a function with the following properties:*

1. *v is strictly concave,*
2. *$\mathcal{M} \subseteq \{ v(z, \cdot) : z \in Z \}$,*
3. *$v(\cdot, x)$ is w-uniformly continuous for all $x \in \mathcal{C}$,*
4. *Either*

   (a) $v(z, \cdot)$ is w-uniformly continuous for all $z \in Z$,

   or

   (b) $v(z, \cdot)$ is Gâteaux-differentiable with continuous derivative $v'(z, \cdot)$ for all $z \in Z$.

According to Assumption 5.30, we are concerned with the maximisation problem

$$\forall z \in Z \qquad \sup_{r \in \mathcal{C}} v(z, r). \tag{5.7}$$

Proposition 5.31. *The set of solutions of* (5.7) *is a closed convex set for all $z \in Z$, which is possibly empty.*

Proof. Let us assume that the supremum in (5.7) is denoted by α. We notice that for every $z \in Z$, the set of solutions of (5.7) is:

$$\{t \in \mathcal{C} : v(z, t) = \alpha\}.$$

Instead, we consider the set $A = \{t \in \mathcal{C} : -v(z, t) \leq -\alpha\}$ and since for every $z \in Z$, $v(z, \cdot)$ is a concave function, then $-v(z, \cdot)$ is convex, therefore A is a convex set of $\mathcal{C}$ (see page 8 in Ekeland and Temam [20]). Since for all $z \in Z$, $v(z, \cdot)$ is Gâteaux-differentiable with continuous derivative $v'(z, \cdot)$, $v(z, \cdot)$ is continuous. Hence, due to the definition of the continuous functions, the set of solutions of (5.7) is also closed. □

Proposition 5.32. *According to Assumption 5.30, problem* (5.7) *has a solution for each $z \in Z$ and this solution is unique.*

Proof. (First proof). In the first part of the proof, we work in the nonstandard universe as in Section 5.1.2. Since $\mathcal{C} \subset W$ is weakly compact, it is bounded in norm. For proving this, let arbitrary $\phi \in W'$. Suppose $\sup_{x\in\mathcal{C}}|\phi(x)|$ is infinite. By the Transfer Principle, there is $c \in {}^*\mathcal{C}$ such that ${}^*\phi(c)$ is infinite. But this is impossible, as Theorem 5.23 shows that there is an $a \in \mathcal{C}$ such that $a \approx_w c$. Therefore $\mathcal{C}$ is bounded in norm, because (see proposition 2.24 in Ng [49]) a subset $\mathcal{C}$ of a normed linear space W is bounded if and only if for all $\phi \in W'$,

$$\sup_{x\in\mathcal{C}}|\phi(x)| < \infty.$$

Since ϕ is arbitrary, the proof is complete. Now let $(t_n)_{n\in\mathbb{N}}$ be a maximising sequence of (5.7), that is, a sequence of elements of $\mathcal{C}$ such that:

$$v(z, t_n) \to \sup_{r\in\mathcal{C}} v(z, r) = \alpha.$$

The set $\mathcal{C}$ is bounded, then the sequence $(t_n)_{n\in\mathbb{N}}$ is itself bounded, and thus we can extract from $(t_n)_{n\in\mathbb{N}}$ a subsequence $(t_{n_m})_{m\in\mathbb{N}}$, which converges weakly to an element t belonging to $\mathcal{C}$ (since W is a reflexive Banach space and this implies that every bounded sequence admits a weakly converging subsequence).

For every $z \in Z$, $v(z, .)$ is a continuous function on $\mathcal{C}$ in the weak topology of W, and hence

$$v(z, t) = lim_{m\to\infty} v(z, t_{n_m}) = \alpha,$$

which means that t is a solution of (5.7).

If two different solutions t_1 and t_2 exist, then from Proposition 5.31, $\frac{t_1+t_2}{2}$ is also a solution. But $v(z, .)$ is strictly concave for every $z \in Z$ and we have

$$v\left(z, \frac{t_1 + t_2}{2}\right) > \frac{1}{2}\left(v(z, t_1) + v(z, t_2)\right) = \alpha,$$

which is a contradiction. □

Proof. (Second proof). Weak compactness implies closedness in the weak topology and since the weak topology is weaker than the norm topology, every weakly open set is open in norm, and by taking complements, every weakly closed set is closed in norm. Therefore both $\mathcal{C}$ and Z are closed. They are also bounded in norm (according to the first proof).

Since v is continuous[1] and $Z \times \mathcal{C}$ is closed and bounded, $v(Z \times \mathcal{C})$ is also a closed and bounded subset of $\mathbb{R}$. According to the boundedness of $v(Z \times \mathcal{C})$ in $\mathbb{R}$, it has a least upper bound α and since it is closed, it contains α. Therefore problem (5.7) has a solution. Uniqueness holds due to the same argument as above. □

[1]Gâteaux-differentiability with continuous derivative and w-uniformly continuity imply continuity.

Remark 5.33. For all $u \in \mathcal{M}$, u attains its unique global maximum on $\mathcal{C}$.

Definition 5.34. $f : {}^*Z \to {}^*\mathbb{R}$ is said to be *w-S-continuous in* $\bar{z} \in {}^*Z$ if and only if

$$\forall \tilde{z} \in Z \qquad (\tilde{z} \approx_w \bar{z} \Rightarrow f(\tilde{z}) \approx f(\bar{z}))\,.$$

f is *w-S-continuous on* ${}^*Z$ if and only if for all weakly nearstandard[2] $\bar{z} \in {}^*Z$, f is w-S-continuous.

Proposition 5.35. *If* $f : Z \to \mathbb{R}$ *is w-uniformly continuous, then* ${}^*f : {}^*Z \to {}^*\mathbb{R}$ *is w-S-continuous.*

Proof. Consider a weakly nearstandard $\bar{z} \in {}^*Z$. Let $\tilde{z} \in Z$ such that $\tilde{z} \approx_w \bar{z}$. Since f is w-uniformly continuous, it is also w-uniformly continuous in $\tilde{z}$, and therefore

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall z \in Z \quad \big((\forall \phi \in X' \quad \|z - \tilde{z}\|_\phi < \delta) \Rightarrow |f(z) - f(\tilde{z})| < \epsilon \big)\,.$$

The Transfer Principle can be applied to

$$\forall z \in Z \quad \big((\forall \phi \in X' \quad \|z - \tilde{z}\|_\phi < \delta) \Rightarrow |f(z) - f(\tilde{z})| < \epsilon \big)\,.$$

Hence

$$\forall \epsilon \in \mathbb{R}_{>0}\, \exists \delta \in \mathbb{R}_{>0}\, \forall z \in {}^*Z \, \big((\forall \phi \in {}^*X' \, \|z - \tilde{z}\|_\phi < \delta) \Rightarrow |{}^*f(z) - {}^*f(\tilde{z})| < \epsilon \big)\,. \quad (5.8)$$

However, $\forall \phi \in {}^*X'$, $\|\bar{z} - \tilde{z}\|_\phi < \delta$ for all $\delta \in \mathbb{R}_{>0}$. Then $|{}^*f(\bar{z}) - {}^*f(\tilde{z})| < \epsilon$ for all $\epsilon \in \mathbb{R}_{>0}$ by (5.8). Therefore ${}^*f(\bar{z}) \approx {}^*f(\tilde{z})$. □

5.2.2 Representative utility function

We recall from Chapter 3 that an admissible utility function $\varphi : \mathcal{C} \to \mathbb{R}$ is said to be F_σ*-socially acceptable* for $\underline{u}$ if and only if there exists some $\tilde{x} \in \mathcal{C}$ with $\varphi(\tilde{x}) = \sup \varphi$ such that for every $y \in \mathcal{C} \backslash \{\tilde{x}\}$, the coalition of i with $u_i(\tilde{x}) > u_i(y)$ is decisive. An admissible utility function $\varphi : \mathcal{C} \to \mathbb{R}$ is called σ*-representative* of $\underline{P} \in \mathcal{P}^N$ if and only if there exists some $\tilde{x} \in \mathcal{C}$ with $\varphi(\tilde{x}) = \sup \varphi$ and any such $\tilde{x}$ is also $\sigma(\underline{P})$-maximal.

Theorem 5.36. *Suppose Assumptions 5.26 and 5.30 hold and* F_σ *is a non-principal ultrafilter. Then there exists for every* $\underline{u} \in \mathcal{M}^N$ *some* F_σ*-socially acceptable utility function.*

Proof. We prove this theorem in two cases:
Case 1 (If properties 1, 2, 3 and 4(a) from Assumption 5.30 are satisfied). Fix an arbitrary $\underline{u} \in \mathcal{M}^N$ and by Assumption 5.30, let $(z_i)_{i \in N} \in (Z)^N$ be such that $u_i = v(z_i, \cdot)$

[2] Weakly nearstandard: nearstandard w.r.t. the weak topology.

for every $i \in N$. The ultrapower construction can easily be adapted to construct an embedding

$$^* : V\left((X \oplus W) \cup \mathbb{R}\right) \to V(^*(X \ \oplus \ W) \cup {}^*\mathbb{R}),$$

where $^*(X \ \oplus \ W)$ is a Banach space over $^*\mathbb{R}$ satisfying the Extension and Transfer Principles. For the rest of the proof, we work in the resulting nonstandard universe. We have to construct some parameter $\tilde{z}$ such that $v(\tilde{z}, \cdot)$ is F_σ-socially acceptable. Let

$$\bar{z} := [(z_i)_{i \in N}]_{F_\sigma} \in {}^*Z.$$

Since Z is a weakly compact subset of a given separable Banach space, by Theorem 5.23 every element of $^*Z$ is nearstandard w.r.t. the weak topology, and let $\tilde{z} \in Z$ such that $\tilde{z} \approx_w \bar{z}$.

Now consider the map

$$t : Z \longrightarrow \mathcal{C},$$

which assigns to each $z \in Z$ the unique $x = t(z) \in \mathcal{C}$ such that

$$x \in \arg\sup_{r \in \mathcal{C}} v(z, r).$$

Note that the existence and uniqueness of x follow from Remark 5.33. By the Transfer Principle,

$$^*t : {}^*Z \to {}^*\mathcal{C},$$

hence $^*t(\bar{z}) \in {}^*\mathcal{C}$ and since $\mathcal{C}$ is a weakly compact subset of a given reflexive separable Banach space, every element of $^*\mathcal{C}$ is nearstandard w.r.t. the weak topology (see again Theorem 5.23) and therefore $\bar{x} = {}^*t(\bar{z})$ is nearstandard w.r.t. the weak topology. Let $\tilde{x} \approx_w \bar{x}$.

Due to Assumption 5.30 (see properties 3 and 4(a)), v is w-uniformly continuous and hence by Proposition 5.35, $^*v$ is w-S-continuous. Therefore, we have for all $y \in \mathcal{C}$,

$$v(\tilde{z}, y) - v(\tilde{z}, \tilde{x}) \approx {}^*v(\bar{z}, y) - {}^*v(\bar{z}, \tilde{x}) \approx {}^*v(\bar{z}, y) - {}^*v(\bar{z}, \bar{x}).$$

The right-hand side of the last equation is a non-positive hyperreal (since $\bar{x}$ is a global *maximum of $^*v(\bar{z}, \cdot)$), so the standard part is non-positive. But the standard part is exactly $v(\tilde{z}, y) - v(\tilde{z}, \tilde{x})$, so

$$v(\tilde{z}, y) - v(\tilde{z}, \tilde{x}) \leq 0 \quad , \qquad \forall y \in \mathcal{C}.$$

Since we have a unique global maximum (according to Remark 5.33),

$$v(\tilde{z}, y) - v(\tilde{z}, \tilde{x}) < 0 \quad , \qquad \text{for all}\, y \neq \tilde{x}. \tag{5.9}$$

In order to verify that $v(\tilde{z},\cdot)$ is F_σ-socially acceptable, we still need to show that for every $y \in \mathcal{C}\backslash\{\tilde{x}\}$, the set of all $i \in N$ with $u_i(\tilde{x}) > u_i(y)$ is decisive (i.e. $\in F_\sigma$). Define a function f by $f(h) := v(h,\tilde{x}) - v(h,y)$ for all $h \in Z$, from where

$$\begin{aligned}\{i \in N \,:\, u_i(\tilde{x}) > u_i(y)\,\} &= \{i \in N \,:\, v(z_i,\tilde{x}) - v(z_i,y) > 0\,\}\\ &= \{i \in N \,:\, f(z_i) > 0\,\}.\end{aligned} \tag{5.10}$$

Due to the construction of the nonstandard embedding * via the bounded ultrapower (with respect to F_σ) of the superstructure $V\left((X \oplus W) \cup \mathbb{R}\right)$, one has the equivalence[3] $\{i \in N \,:\, f(z_i) > 0\,\} \in F_\sigma \;\Leftrightarrow\; {}^*f(\bar{z}) > 0$ which through equation (5.10) yields

$$\{i \in N \,:\, u_i(\tilde{x}) > u_i(y)\,\} \in F_\sigma \;\Leftrightarrow\; {}^*f(\bar{z}) > 0. \tag{5.11}$$

However, by applying the Transfer Principle to the defining equation for f, we get ${}^*f(h) = {}^*v(h,\tilde{x}) - {}^*v(h,y)$ for all $h \in {}^*Z$, so by the w-S-continuity of ${}^*v$ (see Proposition 5.35), we have

$${}^*f(\bar{z}) = {}^*v(\bar{z},\tilde{x}) - {}^*v(\bar{z},y) \approx {}^*v(\bar{z},\bar{x}) - {}^*v(\bar{z},y).$$

The standard part of the right-hand side is strictly positive (by inequality (5.9)) and therefore $^\circ({}^*f(\bar{z})) > 0$. Hence ${}^*f(\bar{z}) > 0$ and by equivalence (5.11) we have

$$\{i \in N \,:\, u_i(\tilde{x}) > u_i(y)\,\} \in \mathcal{D}.$$

Case 2 (If properties 1, 2, 3 and 4(b) from Assumption 5.30 are satisfied). Fix an arbitrary $\underline{u} \in \mathcal{M}^N$ and by Assumption 5.30, let $(z_i)_{i\in N} \in (Z)^N$ be such that $u_i = v(z_i,\cdot)$ for every $i \in N$. Like in the proof of the previous case, we work in the resulting nonstandard universe with the same embedding. We have to construct some parameter $\tilde{z}$ such that $v(\tilde{z},\cdot)$ is F_σ-socially acceptable. Let

$$\bar{z} := [(z_i)_{i\in N}]_{F_\sigma} \in {}^*Z.$$

Since Z is a weakly compact subset of a given separable Banach space, by Theorem 5.23 every element of ${}^*Z$ is nearstandard w.r.t. the weak topology, and let $\tilde{z} \in Z$ such that $\tilde{z} \approx_w \bar{z}$.

Now consider the map

$$t : Z \longrightarrow \mathcal{C},$$

which assigns to each $z \in Z$ the unique $x = t(z) \in \mathcal{C}$ such that

$$x \in \arg\sup_{r\in\mathcal{C}} v(z,r).$$

[3]The sequence $(f(z_i))_{i\in N}$ is bounded in $V\left((X \oplus W) \cup \mathbb{R}\right)$, since $z_i \in Z \subseteq X$ for all $i \in N$ and $f : Z \to \mathbb{R}$.

Note that the existence and uniqueness of x follow from Remark 5.33. Since $v(\tilde{z},\cdot)$ is strictly concave, by duality of the Proposition 5.4 on page 24 in Ekeland and Temam[4] [20], for all $y \in \mathcal{C}$ with $y \neq x$ we have

$$v(\tilde{z},y) < v(\tilde{z},x) + \left\langle v'(\tilde{z},x), y-x \right\rangle.$$

Due to duality of the Proposition 2.1 on page 36 in Ekeland and Temam[5] [20],

$$\left\langle v'(\tilde{z},x), y-x \right\rangle \leq 0.$$

Hence

$$\forall y \in \mathcal{C} \quad y \neq x \qquad v(\tilde{z},y) - v(\tilde{z},x) < 0. \tag{5.12}$$

This proves that x is the unique maximum of $v(\tilde{z},\cdot)$. Like in the proof of Case 1, in order to verify that $v(\tilde{z},\cdot)$ is F_σ-socially acceptable, we define a function f by

$$f(h) := v(h,x) - v(h,y),$$

for all $h \in Z$. Hence

$$\{i \in N \,:\, u_i(x) > u_i(y)\} = \{i \in N \,:\, f(z_i) > 0\}.$$

Due to the construction of the nonstandard embedding * via the bounded ultrapower (with respect to F_σ) of the superstructure $V\left((X \oplus W) \cup \mathbb{R}\right)$, one has the equivalence

$$\{i \in N \,:\, u_i(x) > u_i(y)\} \in F_\sigma \;\Leftrightarrow\; {}^*f(\bar{z}) > 0. \tag{5.13}$$

However, by applying the Transfer Principle to the defining equation for f, we get ${}^*f(h) = {}^*v(h,x) - {}^*v(h,y)$ for all $h \in {}^*Z$, so

$${}^*f(\bar{z}) = {}^*v(\bar{z},x) - {}^*v(\bar{z},y).$$

By the w-S-continuity of ${}^*v(\cdot,x)$ and inequality (5.12), the standard part of the right-hand side is strictly positive and therefore ${}^\circ({}^*f(\bar{z})) > 0$. Hence ${}^*f(\bar{z}) > 0$ and by

[4] Proposition 5.4 in Ekeland and Temam: Suppose Z and $\mathcal{C}$ are as in Assumptions 5.26 and 5.30. If $v : Z \times \mathcal{C} \to \mathbb{R}$ is Gâteaux-differentiable, then for all $z \in Z$, $v(z,\cdot)$ is strictly convex if and only if

$$\forall y, x \in \mathcal{C}, \quad y \neq x \qquad v(z,y) > v(z,x) + \left\langle v'(z,x), y-x \right\rangle.$$

[5] Duality of the Proposition 2.1 in Ekeland and Temam: We assume that v satisfies Assumption 5.30. Then $x \in \mathcal{C}$ is a solution of the maximisation problem 5.7 if and only if

$$\forall y \in \mathcal{C} \quad \forall z \in Z \qquad \left\langle v'(z,x), y-x \right\rangle \leq 0.$$

equivalence (5.13) we have

$$\{i \in N \, : \, u_i(x) > u_i(y)\} \in F_\sigma.$$

□

Theorem 5.37. *Suppose σ satisfies all Arrovian rationality axioms (Axioms 3.10, 3.11 and 3.12), and F_σ is a non-principal ultrafilter. If, in addition, $\mathcal{M}$ satisfies Assumption 5.30, then there exists for every $\underline{u} \in \mathcal{M}^N$ some admissible utility function which is σ-representative of the preference profile $\underline{P}^{\underline{u}}$ induced by $\underline{u}$.*

Proof. The proof is similar to the second part of the proof of Theorem 3.24 in Chapter 3.

□

Chapter 6

Extension: Kirman-Sondermann correspondence for vote abstention

In this chapter we study an aggregation problem with the model theoretic approach by means of *generalised ultraproducts.* Typically in the aggregation problems, we use some ultrafilters on the set of individuals to model the set of decisive coalitions, and as we have shown in Chapter 3 (Section 3.2), there is a one to one correspondence between the Arrow-rational aggregator and the set of decisive coalitions (which is an ultrafilter) induced by this aggregator. Herzberg and Eckert in [32] studied generalisation of the Kirman-Sondermann correspondence. Since some voters might abstain from voting, in the profile some coordinates might be empty. At the moment, the vote abstention situation is not satisfactorily covered by the generalised Kirman-Sondermann correspondence in [32], therefore we generalise this correspondence by using the generalised ultraproduct defined in Makkai [41] in which empty models in some coordinates do not necessarily make the generalised ultraproduct also empty.

6.1 Arrow-rational aggregators

Fix a first-order language $\mathcal{L}$, consisting of constant symbols $\dot{a}$ for each element a in a given set A and countably many n-ary relation symbols $\dot{R}_n$, for every $n \in \mathbb{N}$. Let $\mathcal{S}$ be the set of the atomic $\mathcal{L}$-formulae, and let $\mathcal{I}$ be the *Boolean closure* of $\mathcal{S}$, i.e. the closure of $\mathcal{S}$ under the logical connectives $\dot{\neg}, \dot{\wedge}, \dot{\vee}$. Fix a consistent set T of universal $\mathcal{L}$-sentences[1].

Recall from Chapter 2, that, an $\mathcal{L}$-structure is a pair $\mathfrak{A} = \langle A, (R_n)_{n\in\mathbb{N}}\rangle$ where $R_n \subseteq A^n$. In this chapter, we assume that all $\mathcal{L}$-structures have the domain A, and the constant symbol $\dot{a}$ is always interpreted by a, for all $a \in A$.

[1]A sentence is *universal* if it has the form $(\dot{\forall}\dot{v}_{k_0})\dots(\dot{\forall}\dot{v}_{k_{n-1}})\phi$ for some formula ϕ that does not contain any quantifiers.

Again, due to Chapter 2, an $\mathcal{L}$-structure $\mathfrak{A}$ is a *model*[2] of the theory T if $\mathfrak{A} \models \phi$ for all $\phi \in T$. Let Ω be the collection of models M of T which contains at least two non-isomorphic models[3] with domain A. Sometimes, abusing notation, we will use M for elements in $\Omega \cup \{\emptyset\}$. We let $\dot{R}_n, \ldots, \dot{a}, \ldots$ denote the symbols in the language $\mathcal{L}$ and let $R_n^M, \ldots, a^M, \ldots$ denote the corresponding semantic object in the model M. We use $|M|$ to denote the domain of M and sometimes we use M to denote the domain directly. As usual, for any M and λ, we write $M \models \lambda$ to indicate that λ is true of M.

Fix a non-empty set I, which we will think of as the set of *individuals*, and the subsets of I are called *coalitions*. Elements $\underline{M} \in (\Omega \cup \{\emptyset\})^I$ are the *profiles*. This is toward the generalisation of [32] to the case of vote abstention. For any such profile, and any $\lambda \in \mathcal{I}$, the *coalition supporting* λ *given* $\underline{M}$ is the set $C(\underline{M}, \lambda) := \{i \in I \,:\, M_i \models \lambda\}$.

An *aggregator* is a partial map $f : (\Omega \cup \{\emptyset\})^I \to \Omega \cup \{\emptyset\}$. The domain of f is denoted by *dom*(f). We impose the following axioms for the aggregator:

Axiom 6.1. (Universal Domain). *$f : (\Omega \cup \{\emptyset\})^I \to \Omega \cup \{\emptyset\}$ is a total map, i.e.,*

$$dom(f) = (\Omega \cup \{\emptyset\})^I.$$

Axiom 6.2. (Generalized Pareto Principle). *For all $\underline{M} \in dom(f)$ and all $\lambda \in \mathcal{I}$,*

$$\text{if } f(\underline{M}) \models \lambda, \text{ then } C(\underline{M}, \lambda) \neq \emptyset.$$

Axiom 6.3. (Generalised Systematicity). *For all $\underline{M}, \underline{N} \in dom(f)$ and all $\lambda, \mu \in \mathcal{I}$,*

$$\text{if } C(\underline{M}, \lambda) = C(\underline{N}, \mu), \text{ then } f(\underline{M}) \models \lambda \Leftrightarrow f(\underline{N}) \models \mu.$$

Definition 6.4. An aggregator f is *Arrow-rational* if it satisfies the Axioms 6.1, 6.2 and 6.3. The collection of Arrow-rational aggregators is denoted by $\mathcal{AR}$.

Remark 6.5. Note that Systematicity is the conjunction of Independence of Irrelevent Alternatives (IIA)[4] and Neutrality[5].

Definition 6.6. (Decisive coalition). For any aggregator f, a coalition $C \subseteq I$ is f-*decisive* if, for all $\lambda \in \mathcal{I}$ and all $\underline{M} \in \text{dom}(f)$,

$$\text{if } C = C(\underline{M}, \lambda), \text{ then } f(\underline{M}) \models \lambda.$$

[2]That is, if all sentences of the theory hold true in $\mathfrak{A}$, with the usual Tarski's definition of truth 2.33. For instance, if $\mathfrak{A} = \langle A, (R_n)_{n \in \mathbb{N}} \rangle$, then for all $a_0, a_1, \ldots, a_{n-1} \in A$, one has

$$\mathfrak{A} \models \dot{R}_n \dot{(}\dot{a}_0, \dot{a}_1, \ldots, \dot{a}_{n-1}\dot{)} \Leftrightarrow (a_0, a_1, \ldots, a_{n-1}) \in R_n.$$

[3]This is a crucial point in the proof of Lemma 6.8.

[4]Independence of Irrelevent Alternatives (IIA): For all $\lambda \in \mathcal{I}$ and all profiles $\underline{M}, \underline{N} \in \text{dom}(f)$, if $C(\underline{M}, \lambda) = C(\underline{N}, \lambda)$, then $f(\underline{M}) \models \lambda \Leftrightarrow f(\underline{N}) \models \lambda$.

[5]Neutrality: For all $\lambda, \mu \in \mathcal{I}$ and all profile $\underline{M} \in \text{dom}(f)$, if $C(\underline{M}, \lambda) = C(\underline{M}, \mu)$, then $f(\underline{M}) \models \lambda \Leftrightarrow f(\underline{M}) \models \mu$.

Let $\mathcal{D}_f$ denote the set of the f-decisive coalitions, i.e.

$$\mathcal{D}_f := \{C \subseteq I \,:\, \text{for all } \lambda \in \mathcal{I} \text{ for all } \underline{M} \in \text{dom}(f),\ \text{if } C = C(\underline{M}, \lambda) \text{ then } f(\underline{M}) \models \lambda\}.$$

Lemma 6.7. *If f satisfies Axiom 6.3, then for all $\underline{M} \in \text{dom}(f)$ and $\lambda \in \mathcal{I}$, we have:*

$$C(\underline{M}, \lambda) \in \mathcal{D}_f \Leftrightarrow f(\underline{M}) \models \lambda.$$

Proof. For all $\underline{M} \in dom(f)$ and $\lambda \in \mathcal{I}$:
($\Rightarrow$): If $C(\underline{M}, \lambda) \in \mathcal{D}_f$ then $\forall \lambda' \in \mathcal{I}$ and $\forall \underline{M}' \in dom(f)$,

$$C(\underline{M}, \lambda) = C(\underline{M}', \lambda') \quad \text{implies} \quad f(\underline{M}') \models \lambda'.$$

By taking $\lambda' = \lambda$ and $\underline{M}' = \underline{M}$, we get $f(\underline{M}) \models \lambda$.
($\Leftarrow$): If $f(\underline{M}) \models \lambda$, in order to show $C(\underline{M}, \lambda) \in \mathcal{D}_f$, it suffices to show that for all $\underline{M}' \in \text{dom}(f)$ and for all $\lambda' \in \mathcal{I}$,

$$\text{if} \quad C(\underline{M}, \lambda) = C(\underline{M}', \lambda'), \quad \text{then} \quad f(\underline{M}') \models \lambda'.$$

Now given $\underline{M}', \lambda'$ such that $C(\underline{M}, \lambda) = C(\underline{M}', \lambda')$. By Axiom 6.3,

$$f(\underline{M}) \models \lambda \quad \text{iff} \quad f(\underline{M}') \models \lambda',$$

and since $f(\underline{M}) \models \lambda$, we have $f(\underline{M}') \models \lambda'$. □

Lemma 6.8. *Let $f \in \mathcal{AR}$, then $\mathcal{D}_f$ is an ultrafilter over I.*

Proof. We have to verify the ultrafilter properties for $\mathcal{D}_f$. First, since f satisfies Axiom 6.2, $\mathcal{D}_f$ cannot contain $\emptyset$.

Secondly, we have to show that if $C \in \mathcal{D}_f$ and $C \subseteq C' \subseteq I$, then $C' \in \mathcal{D}_f$. According to our requirement that Ω contains at least two non-isomorphic models of domain A, we take two of them N and N' such that for a relational atomic formula ϕ, $N \models \phi$, $N' \not\models \phi$ and $\emptyset \not\models \phi$. Now take an $a \in A$. It is easy to see that $N \models \dot{a} \doteq \dot{a}$, $N' \models \dot{a} \doteq \dot{a}$ and $\emptyset \not\models \dot{a} \doteq \dot{a}$ [6]. We define the profile $\underline{M}$ as follows:

$$M_i = \begin{cases} N' & \text{if } i \in C \\ N & \text{if } i \in C' \backslash C \\ \emptyset & \text{if } i \in I \backslash C' \end{cases}$$

[6] The $\emptyset$ is in fact the empty model, and all the atomic sentences are false in the empty model. Hence $\dot{a} \doteq \dot{a}$ is not true in it, but $\neg(\dot{a} \doteq \dot{a})$ is true.

Then $C(\underline{M}, \dot{a} \doteq \dot{a} \mathbin{\dot\wedge} \dot\neg\phi) = C$ and $C(\underline{M}, \dot{a} \doteq \dot{a}) = C'$. Therefore

$$\begin{aligned} C(\underline{M}, \dot{a} \doteq \dot{a} \mathbin{\dot\wedge} \dot\neg\phi) = C \in \mathcal{D}_f &\Rightarrow f(\underline{M}) \models \dot{a} \doteq \dot{a} \mathbin{\dot\wedge} \dot\neg\phi && \text{(by Lemma 6.7)} \\ &\Rightarrow f(\underline{M}) \models \dot{a} \doteq \dot{a} \\ &\Rightarrow C' = C(\underline{M}, \dot{a} \doteq \dot{a}) \in \mathcal{D}_f. && \text{(by Lemma 6.7)} \end{aligned}$$

Thirdly we have to show that $\mathcal{D}_f$ is closed under intersection. Let $C, C' \in \mathcal{D}_f$. We take two non-isomorphic models of domain A, namely N and N' such that there exists $\phi \in \mathcal{S}$ with $N \models \phi$, $N' \not\models \phi$ and $\emptyset \not\models \phi$. Now take an $a \in A$. It is easy to see that $N \models \dot{a} \doteq \dot{a}$, $N' \models \dot{a} \doteq \dot{a}$ and $\emptyset \not\models \dot{a} \doteq \dot{a}$. We define the profile $\underline{M}$ as follows:

$$M_i = \begin{cases} N' & \text{if } i \in C \cap C' \\ N & \text{if } i \in C \backslash (C \cap C') \\ \emptyset & \text{if } i \in I \backslash C \end{cases}$$

Then $C(\underline{M}, \dot{a} \doteq \dot{a}) = C \in \mathcal{D}_f$ and by Lemma 6.7, $f(\underline{M}) \models \dot{a} \doteq \dot{a}$. Also $C' \subseteq C(\underline{M}, \dot\neg\phi)$, and since $C' \in \mathcal{D}_f$ and $\mathcal{D}_f$ is closed under supersets (as we have already shown), $C(\underline{M}, \dot\neg\phi) \in \mathcal{D}_f$. Hence, by Lemma 6.7, $f(\underline{M}) \models \dot\neg\phi$. Therefore

$$\begin{aligned} f(\underline{M}) \models \dot{a} \doteq \dot{a} \quad \text{and} \quad f(\underline{M}) \models \dot\neg\phi &\Rightarrow f(\underline{M}) \models \dot{a} \doteq \dot{a} \mathbin{\dot\wedge} \dot\neg\phi \\ &\Rightarrow C(\underline{M}, \dot{a} \doteq \dot{a} \mathbin{\dot\wedge} \dot\neg\phi) \in \mathcal{D}_f. \quad \text{(by Lemma 6.7)} \end{aligned}$$

But $C \cap C' = C(\underline{M}, \dot{a} \doteq \dot{a} \mathbin{\dot\wedge} \dot\neg\phi)$ and hence $C \cap C' \in \mathcal{D}_f$.

It only remains to show that for all $C \subseteq I$, $C \in \mathcal{D}_f$ or $I \backslash C \in \mathcal{D}_f$. We take a model $N \in \Omega$ with domain A and $a \in A$. It is easy to see that $N \models \dot{a} \doteq \dot{a}$ and $\emptyset \not\models \dot{a} \doteq \dot{a}$. We define the profile $\underline{M} \in (\Omega \cup \{\emptyset\})^I$ as follows:

$$M_i = \begin{cases} N & \text{if } i \in C \\ \emptyset & \text{if } i \in I \backslash C \end{cases}$$

Therefore

$$C(\underline{M}, \dot{a} \doteq \dot{a}) = C \quad \text{and} \quad C(\underline{M}, \dot\neg\dot{(}\dot{a} \doteq \dot{a}\dot{)}) = I \backslash C.$$

Since either $f(\underline{M}) \models \dot{a} \doteq \dot{a}$ or $f(\underline{M}) \models \dot\neg\dot{(}\dot{a} \doteq \dot{a}\dot{)}$, by Lemma 6.7 we have

$$C \in \mathcal{D}_f \quad \text{or} \quad I \backslash C \in \mathcal{D}_f,$$

which completes the proof of the lemma. □

6.2 Generalised ultraproduct construction

In the present section, we introduce a construction which, for each (ultra)filter $\mathcal{D}$ over I and each profile $\underline{M} \in (\Omega \cup \{\emptyset\})^I$, yields an $\mathcal{L}$-model. This construction amounts to the specialisation of Makkai's ultraproduct construction (Section 1.3 in [41]) from a more general category-theoretic setting to the model-theoretic setting of interest here. In the remainder of this section we fix a set I and an (ultra)filter $\mathcal{D}$ over I.

Definition 6.9. Let $\underline{M} \in (\Omega \cup \{\emptyset\})^I$ be a profile. We define the *generalised union product* as follows:

$$\coprod_{J \in \mathcal{D}} \prod_{j \in J} M_j := \{(s_i)_{i \in J} \ : \ J \in \mathcal{D} \ \text{and} \ s_i \in |M_i| \ \text{for all} \ i \in J\},$$

where $|M_i|$ is the domain of M_i.

Remark 6.10. If $J \in \mathcal{D}$ and $M_i = \emptyset$ for some $i \in J$, then $\prod_{i \in J} M_i = \emptyset$. Hence

$$\coprod_{J \in \mathcal{D}} \prod_{j \in J} M_j = \coprod_{J \in \mathcal{D}_0} \prod_{j \in J} M_j,$$

where $\mathcal{D}_0 := \{J \in \mathcal{D} \ : \ \forall i (i \in J \Rightarrow M_i \neq \emptyset)\}$. Now consider the case when $\mathcal{D}$ is a principal ultrafilter. In this case $\mathcal{D}$ is generated by the singleton $\{i_0\}$ of an individual i_0 identified with the *dictator*, and we claim that

$$\coprod_{J \in \mathcal{D}} \prod_{j \in J} M_j = \emptyset \qquad \text{iff} \qquad M_{i_0} = \emptyset.$$

For proving this claim, suppose that $\coprod_{J \in \mathcal{D}} \prod_{j \in J} M_j = \emptyset$. We want to show that $M_{i_0} = \emptyset$. Since $\coprod_{J \in \mathcal{D}} \prod_{j \in J} M_j = \emptyset$, for all $J \in \mathcal{D}$, we have

$$\prod_{j \in J} M_j = \emptyset. \tag{6.1}$$

Now take $J := \{i_0\} \in \mathcal{D}$, then $\prod_{j \in J} M_j = M_{i_0}$. But due to equation (6.1), for any choice of J, we have $\prod_{j \in J} M_j = \emptyset$, as well as this case that we take $J := \{i_0\}$. Therefore $M_{i_0} = \emptyset$.

Now assume that $M_{i_0} = \emptyset$. We want to show that $\coprod_{J \in \mathcal{D}} \prod_{j \in J} M_j = \emptyset$, which is equivalent to show that for all $J \in \mathcal{D}$, $\prod_{j \in J} M_j = \emptyset$. Fix an arbitrary $J \in \mathcal{D}$. We need to show that $\prod_{j \in J} M_j = \emptyset$ for the choice $J \in \mathcal{D}$. By our assumption on $\mathcal{D}$, it follows that $i_0 \in J$ and $M_{i_0} = \emptyset$, and this implies that $\prod_{j \in J} M_j = \emptyset$.

Definition 6.11. We define the *equivalence relation* $\equiv_{\mathcal{D}}$ on the non-empty generalised union product $\coprod_{J\in\mathcal{D}}\prod_{j\in J} M_j$ such that

$$(s_j)_{j\in J} \equiv_{\mathcal{D}} (t_k)_{k\in K} \quad \text{iff} \quad \{i \in J\cap K \,:\, s_i = t_i\} \in \mathcal{D},$$

where $J, K \in \mathcal{D}$.

For every $J \in \mathcal{D}$, we denote $\underline{s}^J = (s_j)_{j\in J}$ as an element in generalised union product. Therefore the equivalence class of $\underline{s}^J$ with respect to the equivalence relation $\equiv_{\mathcal{D}}$ shall be denoted by $[\underline{s}^J]_{\equiv_{\mathcal{D}}}$. We also denote the set of all such equivalence classes by $(\coprod_{J\in\mathcal{D}}\prod_{j\in J} M_j)/\equiv_{\mathcal{D}}$.

Definition 6.12. Let $\mathcal{D}$ be an ultrafilter on I and $\underline{M}$ be an element of $(\Omega\cup\{\emptyset\})^I$. Then the *generalised ultraproduct* $(\coprod\prod\underline{M})/\mathcal{D}$ of $\underline{M}$ with respect to $\mathcal{D}$ is defined as follows:

- The domain is $(\coprod_{J\in\mathcal{D}}\prod_{j\in J} M_j)/\equiv_{\mathcal{D}}$.
- For any constant symbol $\dot{a}$, it is interpreted as $[\underline{a}^J]_{\equiv_{\mathcal{D}}}$ where $\underline{a}^J = (a_j)_{j\in J}$ such that $J = \{i \in I \,:\, M_i \neq \emptyset\}$, and $a_j = a$ for all $j \in J$.
- For any n-ary relation symbol $\dot{R}_n$, its interpretation are as follows:
  For any $J_0,\dots J_{n-1} \in \mathcal{D}$ and any $\underline{s_k}^{J_k} \in \prod_{j\in J_k} M_j$ for $k = 0,\dots,n-1$,

  $$\left([\underline{s_0}^{J_0}]_{\equiv_{\mathcal{D}}},\dots,[\underline{s_{n-1}}^{J_{n-1}}]_{\equiv_{\mathcal{D}}}\right) \in R_n^{(\coprod\prod\underline{M})/\mathcal{D}}$$

  if and only if

  $$\{i \in J_0\cap\dots\cap J_{n-1} \,:\, (s_{0,i},\dots s_{n-1,i}) \in R_n^{M_i}\} \in \mathcal{D}.$$

Remark 6.13. If the set $\{i \in I \,:\, M_i = \emptyset\}$ is empty, then the definition above is equivalent to the following, simpler one, which is called the *ultraproduct* $\prod\underline{M}/\mathcal{D}$ of $\underline{M}$ with respect to $\mathcal{D}$:

- The domain is $\prod_{i\in I} M_i/\sim_{\mathcal{D}}$, where for any $\underline{s},\underline{t} \in \prod_{i\in I} M_i$ we have:

  $$(s_i)_{i\in I} \sim_{\mathcal{D}} (t_i)_{i\in I} \quad \text{iff} \quad \{i \in I \,:\, s_i = t_i\} \in \mathcal{D}.$$

- For any constant symbol $\dot{a}$, it is interpreted as $[\underline{a}]_{\sim_{\mathcal{D}}}$ where $\underline{a} = (a_i)_{i\in I}$ and $a_i = a$ for all $i \in I$.
- For any n-ary relation symbol $\dot{R}_n$, its interpretation are as follows:
  For any $\underline{s_k} \in \prod_{i\in I} M_i$ and for $k = 0,\dots,n-1$,

  $$\left([\underline{s_0}]_{\sim_{\mathcal{D}}},\dots,[\underline{s_{n-1}}]_{\equiv_{\mathcal{D}}}\right) \in R_n^{\prod\underline{M}/\mathcal{D}}$$

if and only if
$$\{i \in I : (s_{0,i}, \ldots s_{n-1,i}) \in R_n^{M_i}\} \in \mathcal{D}.$$

Proof. Clearly, for all $(y_i)_{i\in I}$ and $(y'_i)_{i\in I}$,

$$(y_i)_{i\in I} \equiv_{\mathcal{D}} (y'_i)_{i\in I} \quad \text{iff} \quad \{i \in I : y_i = y'_i\} \in \mathcal{D} \quad \text{iff} \quad (y_i)_{i\in I} \sim_{\mathcal{D}} (y'_i)_{i\in I}.$$

Moreover, for every $K \in \mathcal{D}$ and for every $(t_k)_{k\in K}$ there exists some $(y_i)_{i\in I}$ such that $(t_k)_{k\in K} \equiv_{\mathcal{D}} (y_i)_{i\in I}$: indeed let

$$y_i = \begin{cases} t_i & \text{if } i \in K \\ \text{any } y \in M_i \neq \emptyset & \text{otherwise.} \end{cases}$$

This definition guarantees that $\{i \in I \cap K = K : y_i = t_i\} = K \in \mathcal{D}$, and hence $(t_k)_{k\in K} \equiv_{\mathcal{D}} (y_i)_{i\in I}$. From the facts above, it follows that the assignment $[(t_k)_{k\in K}]_{\equiv_{\mathcal{D}}} \mapsto [(y_i)_{i\in I}]_{\sim_{\mathcal{D}}}$ is well defined and has an inverse, given by the assignment $[(y_i)_{i\in I}]_{\sim_{\mathcal{D}}} \mapsto [(y_i)_{i\in I}]_{\equiv_{\mathcal{D}}}$. Therefore the domain of the generalised ultraproduct is equal to

$$\left\{[(s_i)_{i\in I}]_{\equiv_{\mathcal{D}}} : (s_i)_{i\in I} \in \prod_{i\in I} M_i\right\}.$$

For any constant symbol a, its interpretation in generalised ultraproduct is $[(a_i)_{i\in I}]_{\equiv_{\mathcal{D}}}$, since $\{i \in I : M_i \neq \emptyset\} = I$, and its interpretation in ultraproduct is $[(a_i)_{i\in I}]_{\sim_{\mathcal{D}}}$ which corresponds to each other.

For any relation symbol R_n, notice that the domain of the generalised ultraproduct is

$$\left\{[(s_i)_{i\in I}]_{\equiv_{\mathcal{D}}} : (s_i)_{i\in I} \in \prod_{i\in I} M_i\right\},$$

therefore for any $\underline{s_0} = (s_{0,i})_{i\in I}, \ldots, \underline{s_{n-1}} = (s_{n-1,i})_{i\in I} \in \prod_{i\in I} M_i$ we have:

$$\begin{aligned}\left([\underline{s_0}]_{\equiv_{\mathcal{D}}}, \ldots, [\underline{s_{n-1}}]_{\equiv_{\mathcal{D}}}\right) \in R_n^{(\coprod\prod \underline{M})/\mathcal{D}} &\Leftrightarrow \{i \in I \cap \ldots \cap I : (s_{0,i}, \ldots s_{n-1,i}) \in R_n^{M_i}\} \in \mathcal{D} \\ &\Leftrightarrow \left([\underline{s_0}]_{\sim_{\mathcal{D}}}, \ldots, [\underline{s_{n-1}}]_{\sim_{\mathcal{D}}}\right) \in R_n^{(\prod \underline{M})/\mathcal{D}},\end{aligned}$$

which completes the proof of the remark. □

However, if $\emptyset \neq \{i \in I : M_i = \emptyset\}$, then $\prod_{i\in I} M_i/\sim_{\mathcal{D}} = \emptyset$, but the generalised ultraproduct as defined above does not need to be empty.

Theorem 6.14. (Generalised Łoś's Theorem). *If $\lambda(x_0, \ldots, x_{n-1})$ is a formula with n free variables, then for any profile $\underline{M} \in (\Omega \cup \{\emptyset\})^I$, we have:*

$$(\coprod\prod \underline{M})/\mathcal{D} \models \lambda\left([\underline{s_0}^{J_0}]_{\equiv_{\mathcal{D}}}, \ldots, [\underline{s_{n-1}}^{J_{n-1}}]_{\equiv_{\mathcal{D}}}\right)$$

if and only if

$$\{i \in J_0 \cap \ldots \cap J_{n-1} \mid M_i \models \lambda(s_{0,i}, \ldots s_{n-1,i})\} \in \mathcal{D}.$$

Proof. See Makkai [41] page 238. □

6.3 Generalised Kirman-Sondermann correspondence

The present section is aimed at introducing the generalised Kirman-Sondermann correspondence and characterising Arrow-rational aggregators in terms of the generalised ultraproduct construction introduced in the previous section.

Definition 6.15. Let $\mathcal{B}$ be an $\mathcal{L}$-structure with domain B. The *restriction* of $\mathcal{B}$ to $A \subseteq B$ is the $\mathcal{L}$-structure that is obtained by restricting the interpretation of the relation symbols to the domain A. In other words, suppose

$$\left(\coprod\prod \underline{M}\right)/\mathcal{D} = \left\langle \left(\coprod_{J\in\mathcal{D}}\prod_{j\in J} M_j\right) / \equiv_{\mathcal{D}}, \left(R_n^{(\coprod\prod \underline{M})/\mathcal{D}}\right)_{n\in\mathbb{N}} \right\rangle,$$

is a relational structure with $R_n^{(\coprod\prod \underline{M})/\mathcal{D}} \subseteq \left(\left(\coprod_{J\in\mathcal{D}}\prod_{j\in J} M_j\right)/\equiv_{\mathcal{D}}\right)^n$ for each $n \in \mathbb{N}$ and such that there exists a canonical injective map $i : A \to \left(\coprod_{J\in\mathcal{D}}\prod_{j\in J} M_j\right)/\equiv_{\mathcal{D}}$. Then the restriction of $(\coprod\prod \underline{M})/\mathcal{D}$ to A with respect to i is the $\mathcal{L}$-structure

$$\left\langle A, \left(i^{-1}\left[R_n^{(\coprod\prod \underline{M})/\mathcal{D}} \cap i(A)^n\right]\right)_{n\in\mathbb{N}} \right\rangle,$$

and will be denoted by $\mathrm{res}_A(\coprod\prod \underline{M})/\mathcal{D}$.

Corollary 6.16. *Let $\mathcal{D}$ be an ultrafilter on I and $\underline{M} \in (\Omega \cup \{\emptyset\})^I$. Then*

$$res_A(\coprod\prod \underline{M})/\mathcal{D} \models \lambda \Leftrightarrow C(\underline{M}, \lambda) \in \mathcal{D},$$

for all $\lambda \in \mathcal{I}$.

Proof. By the Generalised Łoś's Theorem 6.14, $C(\underline{M}, \lambda) = \{i \in I : M_i \models \lambda\} \in \mathcal{D}$ if and only if $(\coprod\prod \underline{M})/\mathcal{D} \models \lambda$. Since by assumption λ is quantifier-free, the latter condition is equivalent to $\mathrm{res}_A(\coprod\prod \underline{M})/\mathcal{D} \models \lambda$. □

Proposition 6.17. *Let $f \in \mathcal{AR}$, then for all $\underline{M} \in dom(f)$ we have*

$$f(\underline{M}) = res_A\left(\coprod\prod \underline{M}\right)/\mathcal{D}_f,$$

where $\mathcal{D}_f$ is defined in 6.6.

Proof. By Lemma 6.8, $\mathcal{D}_f$ is an ultrafilter, and therefore $f(\underline{M})$ =$\text{res}_A(\coprod\prod \underline{M})/\mathcal{D}_f$ is well-defined for all $\underline{M} \in \text{dom}(f)$. Now we fix an arbitrary $\underline{M} \in \text{dom}(f)$ and $\lambda \in \mathcal{I}$. By using Corollary 6.16 and Lemma 6.7 we get

$$\text{res}_A\left(\coprod\prod \underline{M}\right)/\mathcal{D}_f \models \lambda \Leftrightarrow C(\underline{M}, \lambda) \in \mathcal{D} \Leftrightarrow f(\underline{M}) \models \lambda,$$

therefore

$$\text{res}_A\left(\coprod\prod \underline{M}\right)/\mathcal{D}_f \models \lambda \Leftrightarrow f(\underline{M}) \models \lambda.$$

Since $\lambda \in \mathcal{I}$ was arbitrary, we claim that $\text{res}_A\left(\coprod\prod \underline{M}\right)/\mathcal{D}_f = f(\underline{M})$.

Now it only remains to prove our claim which says that for every $M_1, M_2 \in \Omega \cup \{\emptyset\}$ if

$$\forall\lambda \in \mathcal{I}\ (M_1 \models \lambda \Leftrightarrow M_2 \models \lambda)$$

then $M_1 = M_2$. For the proof we consider three cases:

- If both M_1 and M_2 are empty models, then it is trivial.
- If $M_1 = \emptyset$ and $M_2 \neq \emptyset$, then consider the atomic formula $\dot{a} \doteq \dot{a}$. Hence $M_1 \not\models \dot{a} \doteq \dot{a}$ and $M_2 \models \dot{a} \doteq \dot{a}$. This is a contradiction with our assumption

$$\forall\lambda \in \mathcal{I}\ (M_1 \models \lambda \Leftrightarrow M_2 \models \lambda).$$

- If both M_1 and M_2 are non-empty models, then $|M_1| = |M_2| = A$. If $M_1 = M_2$ then it is trivial. If $M_1 \neq M_2$, then there exists an n-ary relation symbol $\dot{R}_n$ such that $R_n^{M_1} \neq R_n^{M_2}$. Since M_1 and M_2 have the same domain A, both $R_n^{M_1}$ and $R_n^{M_2}$ are subsets of A^n. Hence there exists some $(a_0, \ldots, a_{n-1}) \in A^n$ such that $(a_0, \ldots, a_{n-1}) \in R_n^{M_1}$ but $(a_0, \ldots, a_{n-1}) \notin R_n^{M_2}$ or $(a_0, \ldots, a_{n-1}) \notin R_n^{M_1}$ but $(a_0, \ldots, a_{n-1}) \in R_n^{M_2}$. In both cases

$$(a_0, \ldots, a_{n-1}) \in R_n^{M_1} \not\Leftrightarrow (a_0, \ldots, a_{n-1}) \notin R_n^{M_2}.$$

Therefore

$$M_1 \models \dot{R}_n\dot{(}\dot{a}_0, \ldots, \dot{a}_{n-1}\dot{)} \not\Leftrightarrow M_2 \models \dot{R}_n\dot{(}\dot{a}_0, \ldots, \dot{a}_{n-1}\dot{)},$$

which is a contradiction with our assumption

$$\forall\lambda \in \mathcal{I}\ (M_1 \models \lambda \Leftrightarrow M_2 \models \lambda),$$

and this completes the proof of the lemma.

□

Proposition 6.18. *Let $\mathcal{D}$ be an ultrafilter and suppose the aggregator $f : (\Omega \cup \{\emptyset\})^I \to \Omega \cup \{\emptyset\}$ is defined by*

$$f_{\mathcal{D}}(\underline{M}) = res_A \left(\coprod \prod \underline{M} \right) / \mathcal{D}.$$

Then $f_{\mathcal{D}} \in \mathcal{AR}$.

Proof. Since $\underline{M}$ belongs to $(\Omega \cup \{\emptyset\})^I$, the generalised ultraproduct $(\coprod \prod \underline{M}) / \mathcal{D}$ and its restriction to A are well-defined (notice that $\text{res}_A (\coprod \prod \underline{M}) / \mathcal{D}$ might be an empty model). By using the Generalised Łoś's Theorem 6.14, the generalised ultraproduct $(\coprod \prod \underline{M}) / \mathcal{D}$ is a model of T, and since T is universal, $\text{res}_A (\coprod \prod \underline{M}) / \mathcal{D} \models T$. Therefore $\text{res}_A (\coprod \prod \underline{M}) / \mathcal{D} \in \Omega \cup \{\emptyset\}$. Then $f_{\mathcal{D}}$ is well-defined on $(\Omega \cup \{\emptyset\})^I$ and so it is satisfied Axiom 6.1.

For proving Axiom 6.2, let $\underline{M} \in (\Omega \cup \{\emptyset\})^I$ and $\lambda \in \mathcal{I}$ such that $f_{\mathcal{D}}(\underline{M}) \models \lambda$, that is $\text{res}_A (\coprod \prod \underline{M}) / \mathcal{D} \models \lambda$. Then by using Corollary 6.16, we have $C(\underline{M}, \lambda) \in \mathcal{D}$, hence $C(\underline{M}, \lambda) \neq \emptyset$, since $\mathcal{D}$ is an ultrafilter.

For proving Axiom 6.3, suppose that for all $\underline{M}, \underline{M}' \in (\Omega \cup \{\emptyset\})^I$ and all $\lambda, \lambda' \in \mathcal{I}$ we have $C(\underline{M}, \lambda) = C(\underline{M}', \lambda')$. By using Corollary 6.16 we get

$$f_{\mathcal{D}}(\underline{M}) \models \lambda \Leftrightarrow C(\underline{M}, \lambda) \in \mathcal{D} \Leftrightarrow C(\underline{M}', \lambda') \in \mathcal{D} \Leftrightarrow f_{\mathcal{D}}(\underline{M}') \models \lambda.$$

Hence

$$f_{\mathcal{D}}(\underline{M}) \models \lambda \Leftrightarrow f_{\mathcal{D}}(\underline{M}') \models \lambda'.$$

This completes the proof of the lemma. □

The more general definition is suited to cater for those situations in aggregation theory in which some voters might abstain from voting, thus giving rise to profiles $(M_i)_{i\in I}$ in which some coordinates might be empty. At the moment, the vote abstention situation is not satisfactorily covered by the generalised Kirman-Sondermann correspondence (see Theorem 3.10 in Herzberg and Eckert [32]), because it is not reasonable to assume that the abstention of any single voter would force the outcome of an $\mathcal{AR}$-aggregator $f((M_i)_{i\in I})$ to be empty. However, by adopting the Definition 6.12, Theorem 3.10 in [32] can be generalised to the following:

Theorem 6.19. (Generalised Kirman-Sondermann Correspondence). *There is a bijection between $\mathcal{AR}$ and the set βI of the ultrafilters over I, given by $\Lambda : \mathcal{AR} \to \beta I$, $f \mapsto \mathcal{D}_f$. Its inverse is given by $\Phi : \beta I \to \mathcal{AR}$, $\mathcal{D} \mapsto f_{\mathcal{D}}$, where $f_{\mathcal{D}}$ is the aggregator defined by the Proposition 6.18.*

Proof. For all $f \in \mathcal{AR}$, Lemma 6.8 implies that $\Lambda(f) = \mathcal{D}_f$ is an ultrafilter. For every $f \in \mathcal{AR}$, we have

$$\begin{aligned} \Phi(\Lambda(f)) &= \Phi(\mathcal{D}_f) && \text{(by the definition of } \Lambda) \\ &= f_{\mathcal{D}_f}, && \text{(by the definition of } \Phi) \end{aligned}$$

and for every $\mathcal{D} \in \beta I$ one can write

$$\begin{aligned} \Lambda(\Phi(\mathcal{D})) &= \Lambda(f_{\mathcal{D}}) && \text{(by the definition of } \Phi) \\ &= \mathcal{D}_{f_{\mathcal{D}}}. && \text{(by the definition of } \Lambda) \end{aligned}$$

To show that Φ is the inverse of Λ, it suffices to show that $f_{\mathcal{D}_f} = f$ and $\mathcal{D}_{f_{\mathcal{D}}} = \mathcal{D}$. So for all $\underline{M} \in (\Omega \cup \{\emptyset\})^I$, we have

$$\begin{aligned} f_{\mathcal{D}_f}(\underline{M}) &= \mathrm{res}_A(\coprod \prod \underline{M})/\mathcal{D}_f && \text{(by Proposition 6.18)} \\ &= f(\underline{M}). && \text{(by Proposition 6.17)} \end{aligned}$$

To show $\mathcal{D}_{f_{\mathcal{D}}} = \mathcal{D}$ it suffices to show that for all $X \subseteq I$,

$$X \in \mathcal{D}_{f_{\mathcal{D}}} \Leftrightarrow X \in \mathcal{D}.$$

Given $X \subseteq I$, consider two models $M, M' \in \Omega \cup \{\emptyset\}$ and $\lambda \in \mathcal{I}$ such that $M \models \lambda$ and $M' \not\models \lambda$, and define the profile $\underline{M} \in (\Omega \cup \{\emptyset\})^I$ such that

$$M_i = \begin{cases} M & \text{if } i \in X \\ M' & \text{if } i \notin X \end{cases}$$

then $C(\underline{M}, \lambda) = X$. Now we have

$$\begin{aligned} C(\underline{M}, \lambda) \in \mathcal{D}_{f_{\mathcal{D}}} &\Leftrightarrow f_{\mathcal{D}}(\underline{M}) \models \lambda && \text{(by Lemma 6.7)} \\ &\Leftrightarrow \mathrm{res}_A(\coprod \prod \underline{M})/\mathcal{D} \models \lambda && \text{(by Proposition 6.18)} \\ &\Leftrightarrow C(\underline{M}, \lambda) \in \mathcal{D}. && \text{(by Corollary 6.16)} \end{aligned} \tag{6.2}$$

So

$$X \in \mathcal{D}_{f_{\mathcal{D}}} \Leftrightarrow X \in \mathcal{D}.$$

Thus we have proved $\mathcal{D}_{f_{\mathcal{D}}} = \mathcal{D}$ and this completes the proof of the theorem. □

Chapter 7

Conclusion

The traditional reason for using the assumption of the *representative agent* in mathematical models of macroeconomic theory is that it provides microeconomic foundations for aggregate behaviour. So far there have been few attempts at rigorously justifying this assumption. Our contribution combines Arrovian aggregation theory (on an infinite electorate) with structural assumptions on the individual optimisation problem. We adopt the hypothesis that the social planner's goal is maximising the social welfare function. After aggregating individual preferences, we show that there is a *representative utility function* by proving that the maximiser of this representative utility function is the optimal alternative according to the social preference relation, by using an explicit mathematical construction based on techniques from mathematical logic and analysis. We also provide sufficient conditions for these results to be satisfied in economic applications and afterwards give an example of a possible macroeconomic application.

We establish a new and simpler nonstandard account of the weak topology and study the nonstandard hull with respect to this weak topology. Due to nice assumptions on the class of admissible utility functions, we construct a representative utility function for infinite-dimensional social decision problems under weaker conditions for the set of social alternatives.

Generalised ultraproducts allow us to make a correspondence between Arrow-rational aggregators and ultrafilters on the set of individuals even in the setting of vote abstention. For this purpose, we give the *generalised Kirman-Sondermann correspondence* and characterise Arrow-rational aggregators in terms of the generalised ultraproduct construction. For the generalised ultraproduct, even if there are empty models in some coordinates, the resulting generalised ultraproduct is not necessarily empty. If the set of individuals is finite, then we have a *dictator* which can participate in voting or be absent. When the dictator votes, it is obvious that the aggregation result is non-empty, even other voters are absent. When the dictator is absent, the aggregation result is empty, even if other voters vote. For an infinite number of individuals, if the corresponding ultrafilter is *principal*, then the case is like the finite case. If the corresponding ultrafilter is *nonprincipal*, then the aggregation result is non-empty if and only if there exists a decisive coalition such that each member turns out in the voting (that is, everyone in the dictator group goes to vote).

This setting for vote abstention allows us to prove a strengthened version of Arrow's impossibility theorem which, unlike the standard one, holds also for two candidate elections. More technically, the usual assumption on the existence of three non-isomorphic models of the theory is dropped and replaced by the weaker requirement on the existence of two non-isomorphic models (the proof and more discussion appear in Bedrosian, Palmigiano and Zhao [9]).

Bibliography

[1] G. A. Akerlof. Behavioral macroeconomics and macroeconomic behavior. *American Economic Review*, pages 411–433, 2002.

[2] S. Albeverio, R. Høegh-Krohn, J. E. Fenstad, and T. Lindstrøm. *Nonstandard methods in stochastic analysis and mathematical physics*, volume 122 of *Pure and Applied Mathematics.* Academic Press, Inc., Orlando, FL, 1986.

[3] C. D. Aliprantis, D. J. Brown, and O. Burkinshaw. Valuation and optimality in the overlapping generations model. *International Economic Review*, pages 275–288, 1990.

[4] R. M. Anderson. Nonstandard analysis with applications to economics. In *Handbook of mathematical economics, Vol. IV*, volume 1 of *Handbooks in Econom.*, pages 2145–2208. North-Holland, Amsterdam, 1991.

[5] T. E. Armstrong. Arrow's theorem with restricted coalition algebras. *Journal of Mathematical Economics*, 7(1):55–75, 1980.

[6] T. E. Armstrong. Precisely dictatorial social welfare functions: Erratum and addendum to 'arrows theorem with restricted coalition algebras'. *Journal of Mathematical Economics*, 14(1):57–59, 1985.

[7] K. J. Arrow. *Social Choice and Individual Values.* Cowles Commission Monograph No. 12. John Wiley & Sons, Inc., New York, N. Y.; Chapman & Hall, Ltd., London, 1951.

[8] M. E. Baron and H. J. M. Bos. *History of Mathematics: Origins and Development of the Calculus. Newton and Leibniz/Prepared by Margaret E. Baron and HJM Bos for the [Open University] Course Team. C3.* Open University Press, 1974.

[9] G. Bedrosian, A. Palmigiano, and Z. Zhao. *Generalised ultraproduct and Kirman-Sondermann correspondence for vote abstention.* To appear in Lecture Notes in Computer Science, Springer, 2015.

[10] J. L. Bell and A. B. Slomson. *Models and ultraproducts: An introduction.* North-Holland Publishing Co., Amsterdam-London, 1969.

[11] O. J. Blanchard. Comment on bt mccallum, " new classical macroeconomics: A sympathetic account". *The Scandinavian Journal of Economics*, pages 259–264, 1989.

[12] E. D. Bloch. *The real numbers and real analysis.* Springer Science & Business Media, 2011.

[13] E. v. Böhm-Bawerk. Professor clark's views on the genesis of capital,". *Quarterly Journal of Economics*, 9(2):113–31, 1895.

[14] C. C. Chang and H. J. Keisler. *Model theory.* North-Holland Publishing Co., Amsterdam-London; American Elsevier Publishing Co., Inc., New York, 1973. Studies in Logic and the Foundations of Mathematics, Vol. 73.

[15] J. Child. The early mathematical manuscripts of leibniz, translated from the latin texts, published by carl immanuel gerhardt with critical and historical notes by jm child, 1920.

[16] S. A. Clark. The representative agent model of probabilistic social choice. *Mathematical Social Sciences*, 23(1):45–66, 1992.

[17] E. Deutsch. On vectorial norms and pseudonorms. *Proc. Amer. Math. Soc.*, 28:18–24, 1971.

[18] F. Diener and M. Diener. *Nonstandard analysis in practice.* Springer Science & Business Media, 1995.

[19] E. J. Dijksterhuis. *Archimedes.* Princeton University Press, 2014.

[20] I. Ekeland and R. Temam. *Convex analysis and variational problems.* North-Holland Publishing Co., Amsterdam-Oxford; American Elsevier Publishing Co., Inc., New York, 1976. Translated from the French, Studies in Mathematics and its Applications, Vol. 1.

[21] P. C. Fishburn. Arrow's impossibility theorem: concise proof and infinite voters. *J. Econom. Theory*, 2:103–106, 1970.

[22] P. C. Fishburn. *Utility theory for decision making.* Publications in Operations Research, No. 18. John Wiley & Sons, Inc., New York-London-Sydney, 1970.

[23] M. Gillman. *Advanced modern macroeconomics: analysis and application.* Financial Times Press, 2011.

[24] R. Goldblatt. *Lectures on the hyperreals: an introduction to nonstandard analysis*, volume 188. Springer Science & Business Media, 1998.

[25] R. J. Gordon. Symposium on macroeconomics 1 fresh water, salt water, and other macroeconomic elixirs*. *Economic Record*, 65(2):177–184, 1989.

[26] J. E. Hartley. Retrospectives: The origins of the representative agent. *The Journal of Economic Perspectives*, pages 169–177, 1996.

[27] J. E. Hartley. *The representative agent in macroeconomics.* Routledge, 2002.

[28] K. H. Hennings. *The Austrian Theory of Value and Capital: Studies in the Life and Work of Eugen von Böhm Bawerk.* Cheltenham, UK and Lyme, USA: Edward Elgar, 1997.

[29] C. W. Henson and L. C. Moore, Jr. Nonstandard analysis and the theory of Banach spaces. In *Nonstandard analysis—recent developments (Victoria, B.C., 1980)*, volume 983 of *Lecture Notes in Math.*, pages 27–112. Springer, Berlin, 1983.

[30] F. Herzberg. A definable nonstandard enlargement. *Mathematical Logic Quarterly*, 54(2):167–175, 2008.

[31] F. Herzberg. A representative individual from arrovian aggregation of parametric individual utilities. *Journal of Mathematical Economics*, 46(6):1115–1124, 2010.

[32] F. Herzberg and D. Eckert. The model-theoretic approach to aggregation: Impossibility results for finite and infinite electorates. *Mathematical Social Sciences*, 64(1):41–47, 2012.

[33] R. Jin. Applications of nonstandard analysis in additive number theory. *Bull. Symbolic Logic*, 6(3):331–341, 2000.

[34] A. P. Kirman. Whom or what does the representative individual represent? *The Journal of Economic Perspectives*, pages 117–136, 1992.

[35] A. P. Kirman and D. Sondermann. Arrow's theorem, many agents, and invisible dictators. *Journal of Economic Theory*, 5(2):267–277, 1972.

[36] L. Lauwers and L. Van Liedekerke. Ultraproducts and aggregation. *Journal of Mathematical Economics*, 24(3):217–237, 1995.

[37] L. Ljungqvist and T. J. Sargent. *Recursive macroeconomic theory.* MIT press, 2004.

[38] J. Loś. Quelques remarques, theoremes et problemes sur les classes definissables d'algebres. *Studies in Logic and the Foundations of Mathematics*, 16:98–113, 1955.

[39] R. E. Lucas. Econometric policy evaluation: A critique. In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. Elsevier, 1976.

[40] W. A. J. Luxemburg. A general theory of monads. In *Applications of Model Theory to Algebra, Analysis, and Probability (Inte rnat. Sympos., Pasadena, Calif., 1967)*, pages 18–86. Holt, Rinehart and Winston, New York, 1969.

[41] M. Makkai. Ultraproducts and categorical logic. In *Methods in mathematical logic*, pages 222–309. Springer, 1985.

[42] N. G. Mankiw. *Macroeconomics*. Worth Publishers, 2012.

[43] N. G. Mankiw. *Principles of macroeconomics*. Cengage Learning, 2014.

[44] N. G. Mankiw. *Principles of microeconomics*. Cengage Learning, 2014.

[45] N. G. Mankiw and D. Romer. *New Keynesian Economics: Coordination failures and real rigidities*, volume 2. MIT press, 1991.

[46] D. Marker. *Model theory: an introduction*. Springer Science & Business Media, 2002.

[47] A. Marshall. *Principles of economics: an introductory volume*. Macmillan London, 1920.

[48] T. Negishi. Welfare economics and existence of an equilibrium for a competitive economy. *Metroeconomica*, 12(2-3):92–97, 1960.

[49] S.-A. Ng. *Nonstandard methods in functional analysis*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2010. Lectures and notes.

[50] M. Pivato. Variable-population voting rules. *Journal of Mathematical Economics*, 49(3):210–221, 2013.

[51] A. Robinson. *Non-Standard Analysis: Studies in Logic and the Foundations of Mathematics*. North-Holland Publishing Co., Amsterdam, 1966.

[52] A. Robinson. *Non-standard analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1996. Reprint of the second (1974) edition, With a foreword by Wilhelmus A. J. Luxemburg.

[53] A. Robinson and E. Zakon. A set-theoretical characterization of enlargements. In *Applications of Model Theory to Algebra, Analysis, and Probability (Internat. Sympos., Pasadena, Calif., 1967)*, pages 109–122. Holt, Rinehart and Winston, New York, 1969.

[54] H. Royden and P. Fitzpatrick. *Real analysis*. Prentice Hall, 2010.

[55] S. Salbany and T. Todorov. Nonstandard analysis in topology: nonstandard and standard compactifications. *J. Symbolic Logic*, 65(4):1836–1840, 2000.

[56] T. J. Sargent. *Macroeconomic theory*, volume 2. Academic Press New York, 1987.

[57] J. N. Sharma and A. R. Vasishtha. *Functional analysis*. Krishna Prakashan Media Ltd., 1975.

[58] B. Snowdon and H. R. Vane. *Modern macroeconomics: its origins, development and current state*. Edward Elgar Publishing, 2005.

[59] J. E. Stiglitz. The contributions of the economics of information to twentieth century economics. *Quarterly Journal of economics*, pages 1441–1478, 2000.

[60] A. Tarski. Une contribution à la théorie de la mesure. *Fundamenta Mathematicae*, 1(15):42–50, 1930.

[61] S. Ulam. Concerning functions of sets. *Fundamenta Mathematicae*, 1(14):231–233, 1929.

[62] C. E. Walsh. *Monetary theory and policy*. MIT press, 2010.

[63] E. R. Weintraub. The microfoundations of macroeconomics: a critical survey. *Journal of Economic Literature*, pages 1–23, 1977.

[64] M. Wolff and P. A. Loeb, editors. *Nonstandard analysis for the working mathematician*, volume 510 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 2000.