Universität Bielefeld

# Nonstandard Analysis for $G$-Stochastic Calculus

*Dissertation zur Erlangung des Doktorgrades*
*der Fakultät für Mathematik*
*der Universität Bielefeld*
*vorgelegt von*

**Tolulope Rhoda Fadina**

March 2015

Erstreferent: Prof. Dr. Frank Riedel
Zweitreferent: Prof. Dr. Sergio Albeverio

To my parents

# *Acknowledgements*

I would like to express my sincere gratitude to everyone who has contributed to this work:

I appreciate the support and guidance of my advisors, Prof. Frank Riedel and Jun-Prof. Frederik Herzberg who proposed a challenging, yet very interesting research project and constantly voiced their confidence in my abilities. This thesis wouldn't have been completed without their advice, ideas, and constructive criticism, as well as their devotion to discuss the challenges and progress in my thesis.

I am very grateful to Prof. Shige Peng, Prof. Yan Dolinsky, Prof. Moritz Kaßmann, Dr. Yongsheng Song, Dr. Patrick Beissner and Ludovic Tangpi for helpful comments and suggestions. My appreciation also goes to Prof. Sergio Albeverio and Prof. Thomas Zink.

My gratitude goes to Prof. Michael Röckner, Prof. Philip Blanchard, and all the visiting Professors to the Graduate College (Stochastics and Real World Models) for the insightful lectures. They have indeed broadened my knowledge on current research issues emerging in stochastic analysis and its applications to financial economics, physics and biology.

Financial support received from the German Science Foundation (DFG) through the International Graduate College (IGK) Stochastics and Real World Models (Bielefeld–Beijing) and the Rectorate of Bielefeld University (Bielefeld Young Researchers' Fund) is gratefully acknowledged.

Special thanks to Prof Zhiming Ma and all professors and colleagues at the Graduate University of Chinese Academy of Science (GUCAS) for their kind gesture and support throughout my stay. I would like to thank Rebecca Reischuk, Karin Zelmer and Stephan Merkes for their readiness to always solve my administrative problems. To my colleagues from our Graduate College and Institute of Mathematical Economics, especially Diana, Geghard, Viktor, Mykola, Marina, Jesper, Andrea, Milan and Yuhua, thank you for your support and encouragements.

Finally, my hearty thanks goes to my parents for giving me the best legacy, you are my role model. To my siblings, you are simply the best and I truly appreciate you. Huge thanks to Olumide Olumodimu, for your unconditional love and prayer.

# Contents

# Chapter 1

# General Introduction

This thesis consists of three self-contained essays and a concluding chapter. The summary of the main results of the essays is the following: First, we prove a Donsker result for the $G$-Brownian motion with finite state-space. In the second essay, we give an elementary and more intuitive introduction to nonstandard measure theory and we also provide an alternative construction of the renowned Loeb measure. Following, we develop the basic theory for the hyperfinite $G$-expectation.

Recently, it has become increasingly clear that in addition to randomness that can be captured by probability, financial markets are also sensitive to the so-called *model uncertainty*, in the sense that the probability distribution of randomness is unknown. Thus, it becomes problematic to measure the uncertainty associated to a given financial security in such a market. Faced with this problem, financial economists (researchers and practitioners) need to develop a new model that can capture several sources of uncertainty. The theory of $G$-expectation (also known as sublinear expectation), motivated by coherent risk measures, henceforth referred to as *$G$-stochastic calculus*, introduced by Peng [87, 89], provide a convenient mathematical tool to model uncertainty. The $G$-expectation and its corresponding canonical process, the $G$-Brownian motion, can be seen as the central objects of the $G$-stochastic calculus.

In the **first essay** (see Chapter 2), we refine the discretization of the $G$-expectation by Dolinsky et al. [38], in order to obtain a discretization of the sublinear expectation where the martingale laws are defined on a finite lattice rather than the whole set of reals. Dolinsky et al. [38] introduced a notion of volatility uncertainty on a discrete timeline and defined a sequence of sublinear expectations (discrete

$G$-expectation) on the canonical space of the discrete time paths. By the analogue of Donsker's theorem, the discrete-time sublinear expectation converges to the $G$-expectation on the continuous paths. In their approach, they only discretize the timeline but not the state-space of the canonical process. For certain applications, especially in nonstandard analysis, a discretization of the state space would be necessary. Thus, we develop a modification of the construction by Dolinsky et al. [38] which even ensures that the sublinear expectation operator for the discrete-time canonical process corresponding to this discretization of the state space converges to the $G$-expectation.

Despite the usefulness of the powerful tool of nonstandard analysis, the fact that the language of the theory is based on logic has deterred and limited the number of potential practitioners of nonstandard analysis. Thus, in the **second essay** (see Chapter 3), we give a simplified introduction to nonstandard measure theory that does not presuppose prior acquaintance with mathematical logic. The methodology is presented in terms of sequences, equivalence relations and equivalence classes with respect to binary measures. This procedure is based on Lindstrøm's [73] work. However, our approach is more simplified, in the sense that we construct the extended nonstandard enlargement in measure theoretic language. We also show how the language of logic relates to the mathematical discourse in probability theory. Finally, we provide an alternative construction of the Loeb measure using basic knowledge of real analysis.

The **third essay** (see Chapter 4) provides a mathematical foundation for the application of the powerful tools of nonstandard analysis to $G$-stochastic calculus and also potentially prepares the ground for the application of both nonstandard analysis and $G$-stochastic calculus to financial economics. We apply Robinsonian nonstandard analysis to $G$-stochastic calculus in order to provide an alternative, combinatorially inspired construction of the $G$-expectation. Thus, we prove a lifting theorem for the $G$-expectation. Herein, we use our discretization theorem for the $G$-expectation from Chapter 2, Theorem 2.3.13. Very roughly speaking, we extend the discrete time analogue of the $G$-expectation to a hyperfinite time analogue. Then, we use the characterization of convergence in nonstandard analysis to prove that the hyperfinite discrete-time analogue of the $G$-expectation is infinitely close to the classical $G$-expectation. We hope that this result may eventually become useful for applications in financial economics (especially existence of equilibrium in continuous-time financial markets with volatility uncertainty) and that it provides additional intuition for $G$-stochastic calculus.

## 1.1 Financial markets with model uncertainty

Starting from Bernoulli, through Wiener, and then Kolmogorov, Itô and many more, modern finance received its mathematical foundation. Kolmogorov [63] postulated the fundamental axioms for modern probability theory. The edifice of the theory is built on measure theory introduced by Émile Borel and Henry Lebesque and developed by Radon and Fréchet. A probability space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$. That is, a measurable space $(\Omega, \mathcal{F})$ equipped with a probability measure $\mathbb{P}$. On the one hand, with this space, it is possible to predict future scenarios from current events ($\mathbb{P}$ is used to define null events). On the other hand, if there are many probability measures $\mathbb{P}$ within a given set $\mathcal{P}$ with equal possibilities, then it becomes uncertain which one is the true[1] probability. As such a probability space cannot be defined by any single probability measure, the notion of *expectation* and its non-additive counterpart *capacity* plays a significant role. Kolmogorov's [63] success can also be traced to the idea of conditional expectation, equivalent measures and stochastic processes. These tools make possible the mathematical formulation of financial markets.

Fischer Black and Myron Scholes articulated the ground breaking model for pricing European call and put options in their seminal paper Black and Scholes [18]. The key concepts in their derivation are replication and the no-arbitrage option pricing theory. Ross [96] described the general principle of arbitrage option pricing theory. Accordingly, if a market is arbitrage free, then there exists a probability measure on the future scenarios such that today's worth of the option is the expected discounted payoff[2], Duffie and Protter [40]. This probability measure is called *martingale measure*, see Harrison and Kreps [48]. The uniqueness of a martingale measure translates into market completeness[3]. Harrison and Kreps [48] and Harrison and Pliska [49] introduced the mathematical theory of semimartingales and stochastic integrals needed for adequate financial modeling. Harrison and

[1]True probability gives a perfect statistical description of measurable quantities in a financial market.

[2]This is known today as the first fundamental theorem of asset pricing.

[3]In a complete market, there exists a unique hedging strategy for replication of a contingent claim.

Kreps [48] asserted that both the no-arbitrage and the equilibrium asset prices (in the spirit of Arrow-Debreu equilibrium, Debreu [35]) can be formalized in terms of the state prices [4]. However, this is only possible on a probability space that allows state price density or equivalent martingale measures, in the sense that the state price of a given asset is proportional to the marginal utility consumption of individuals in that state.

A basic problem in general equilibrium theory is to prove the existence of equilibrium in a financial market in which each agent takes the best possible decision and the demand and supply are balanced. The existence of equilibrium in continuous-time Arrow-Debreu economy has been proved in the literature, see Mas-Colell and Richard [81], Dana [33] and Bank and Riedel [8]. The martingale representation theorem links the concept of market completeness with a stochastic spanning condition in Arrow-Debreu economy whose dynamics is driven by Brownian motion. However, in more general security markets, the limitation to trade only a prespecified set of securities results in market incompleteness. Duffie and Huang [39] articulated the idea of Kreps [65] on implementing an Arrow-Debreu allocation into a Radner [91] economy. Such an economy is said to be dynamically complete if agents can acquire all the consumption allocations they could achieve in an Arrow-Debreu market by continuously trading the given set of securities. In continuous-time models with a single agent, a large amount of literature has concentrated on proving the existence of equilibrium in such markets, see for example, Bick [14], Cox et al. [26], Herzberg [53] to mention but a few. All existing literature studies the case where the markets are dynamically complete. In this instance, they establish a standard method of constructing equilibrium in the financial markets. However, the general condition under which equilibrium exists varies. Herzberg [53] provided the foundation for equilibrium analysis in a financial market whose dynamics of the dividends is driven by an exponential Lévy process. He proved the existence of equilibrium in a Radner economy with a single agent and trading using *nonstandard analysis*. Anderson and Raimondo [5] proved the existence of equilibrium in a Radner economy for multiple agents with Brownian information and trading using nonstandard analysis (see Section 1.2 for a detailed discussion). The paper of Anderson and Raimondo [5] was the first in a series of papers to prove the existence of equilibrium for a multi-agent financial market with Brownian information.

[4] The state price for a given state at a particular time dictates how much an investor is ready to part with today in return for an extra payment of a unit in the future state.

### 1.1.1 Model uncertainty

Based on Knight's [62] remark, *risk* can be regarded as randomness that can be captured by probability and *uncertainty*[5] can be regarded as all other forms of randomness. A large amount of literature has focused on uncertainty as it relates to financial markets, see for example Bewley [13], Rigotti and Shannon [93], Epstein and Wang [46], Dana and Riedel [34], Epstein and Schneider [45], Chen and Epstein [22], and Trojani and Vanini [105]. Our economic motivation for studying model uncertainty lies in asset pricing and equilibrium analysis of financial markets in the presence of volatility uncertainty. Ultimately, we aim to prove the existence of equilibrium in a continuous-time financial market with volatility uncertainty and multiple agents. Typically, the required probabilistic setup for most financial economic analysis, assumes that all priors[6] are equivalent (i.e., they agree which events are null). Accordingly, many results in the literature invariably depend on the Girsanov theorem for a change of measure. The drift uncertainty can be reduced to uncertainty in which the equivalent probability measure is the physical probability measure. But in light of Girsanov's theorem, this does not affect pricing, which always occurs with respect to a risk-neutral probability measure (also known as equivalent martingale measure). Thus, drift priors are equivalent (for example, see Chen and Epstein [22] and Cheng and Riedel [23]). However, modeling volatility uncertainty generates a set of priors that are non-equivalent (mutually singular, i.e., they disagree about which events are possible.).

### 1.1.2 Volatility uncertainty

The volatility of a continuous time model is a function of the quadratic variation of the underlying state process. In a sense, this describes the random buffeting power of all sources of risk that influence the financial environment in which the asset price is determined. There is an extensive literature (see for recent examples, Eraker and Shaliastovich [47], Bollerslev et al. [19]) on stochastic time varying volatility models. They often argue that the dynamics of the volatility is driven by complicated structures, for example, the dynamics of volatility of volatility. However, the confidence of modelers in these models is still questionable. Carr

[5]Uncertainty is also known as ambiguity.

[6]A prior is a well known term in economics denoting a probability measure.

and Lee [21] object the assumption of modeling volatility based on a particular parametric process. They remark that:

> *"The problem is particular acute for volatility models because the quantity being modeled is not directly observable. Although an estimate for the initially unobserved state variable can be inferred from market prices of derivative securities, noise in the data generates noise in the estimate, raising doubts that a modeler can correctly select any parametric stochastic process from the menu of consistent alternatives."*

However, the knowledge that the volatility of the state variable lies within a particular confidence interval still remains plausible. One approach to model volatility uncertainty would be to consider several models with the objective of capturing *all* sources of uncertainty that can initiate misspecification of the model parameter.

In the spirit of Epstein and Ji [43], we give an illustration on modeling volatility uncertainty using a trinomial tree.

> In a game we are to pick from a sequence of $n$ independent urns that describe uncertainty and each urn contains 100 balls with three different types: $D$ (Down), $C$ (Constant) or $U$ (Up). The time varies over $\mathbb{T} = \{0, \Delta, 2\Delta, \ldots, n\Delta = T\}$. The dynamic of each urn, that is the state variable $B_t = (B)_{t\in\mathbb{T}}$ with $B_0 = 0$ is given by
>
> $$B_{t\Delta} - B_{(t-1)\Delta} = \begin{cases} +\sqrt{\Delta} & U, \\ 0 & C, \\ -\sqrt{\Delta} & D. \end{cases}$$
>
> - **First scenario**: Here we are told the number of balls of type $U$ equals to the one of type $D$ and there are no balls of type $C$ (uncertainty is weakened here because we have a full hand information on the ball composition in each urn) for each urn. We also know that all the urns are identically composed. The resulting probability measure on the set of increment paths is a product measure whose factors are all identical and given by
>
>   $$\frac{\delta_{-\sqrt{\Delta}} + \delta_{+\sqrt{\Delta}}}{2},$$

where for all $A \subseteq \mathbb{R}$,

$$\delta_x(A) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}.$$

Thus, we have a random walk. By Donsker's theorem, the random walk converges weakly to a Brownian motion in the continuous-time limit as $\Delta$ tends to 0 (see for example, Billingsley [16, Theorem 14.1 and Example 12.3]). Cox et al. [27] uses this approach to derive the Black-Scholes option pricing formula, as the limit of a discrete-time binomial option pricing formula.

- **Second scenario**: We are told the number of balls of type $U$ equals to the one of type $D$ in each urn. But we only know the number of balls of type $C$ is less than 30. Any probability measure on the path of $B$ with this information makes $B$ a martingale. The variance $\sigma_t^2$ of the difference between two paths depends on the number of balls of type $C$ in the urns and its uncertainty is then defined on this range as $\frac{70}{100} = \underline{\sigma}_t^2 \leq \sigma_t^2 \leq \overline{\sigma}_t^2 = 1$. Since the urns are independent, i.e., we cannot predict the future draws from the past draws, the composition of the balls might not be the same. Thus, the value of $\sigma_t$ on the bound $[\underline{\sigma}, \overline{\sigma}]$ might vary. In the continuous-time limit as $\Delta$ goes to 0, the discrete-time trinomial model converges weakly in distribution to a continuous-time model on $[0, T]$ (this type of convergence is discussed below). And the canonical process $B = B_t$ still inherits the martingale property of the discrete-time setting, where $\sigma_t$ lies in the interval $[\underline{\sigma}, \overline{\sigma}]$. To be more detailed about the notion of volatility, we denote the quadratic variation of $B$ as follows:

$$\langle B \rangle_t(\omega) = \lim_{\Delta \to 0} \sum_{s<t} |B_{s\Delta} - B_{(s-1)\Delta}|^2.$$

Then, the volatility $\langle B \rangle_t$ lies in the confidence interval $[\underline{\sigma}_t^2, \overline{\sigma}_t^2]$. It is important to note that volatility uncertainty leads to a set of nonequivalent probability measures. The canonical process $B$ is called the *G-Brownian*.

A very sketchy summary of the used convergence result is the following: Let $\Omega$ be a set of continuous paths on $[0,T]$ that starts at the origin which is zero. Informally, a *sublinear expectation*, $\mathcal{E}(\cdot)$, is a function that is defined on a linear space of random variables that satisfies monotonicity, constant preserving, sub-additivity and positive homogeneity. Consider a sequence of real-valued random variables $(\xi_k)_{k\geq 1}$ such that $\xi_{k+1}$ and $\xi_k$ are identically distributed[7] and $\xi_{k+1}$ is independent[8] from $(\xi_1,\ldots,\xi_k)$ for $k=1,2,\ldots$. We assume that

$$\mathcal{E}[\xi_1]=\mathcal{E}[-\xi_1]=0,\quad \mathcal{E}[-\xi_1^2]=\underline{\sigma}^2 \quad\text{and}\quad \mathcal{E}[\xi_1^2]=\overline{\sigma}^2.$$

We also introduce a sequence of partial sums $Z_n=\sum_{i=1}^n \xi_i$ where $Z_0=0$. By linear interpolation, a continuous process $\widehat{Y}_t$ can be obtained from the sequence of $Z_n$:

$$\widehat{Y}_t=(\lfloor t\rfloor+1-t)Z_{\lfloor t\rfloor}+(t-\lfloor t\rfloor)Z_{\lfloor t\rfloor+1},$$

where $\lfloor y\rfloor$ denotes the greatest integer less than or equal to $y$. For instance, if $Z_0=0$, $Z_1=0.35$, $Z_2=-1.16$, $Z_3=1.58$, $Z_4=0.41$, for $0\leq t<1$, $\widehat{Y}_t = (1-t)Z_0 + tZ_1 = 0$; for $1\leq t<2, \widehat{Y}_t=(2-t)Z_1+(t-1)Z_2=0.35$; for $2\leq t<3$, $\widehat{Y}_t=(3-t)Z_2+(t-2)Z_3=-1.16$, for $3\leq t<4$, $\widehat{Z}_t=(4-t)Z_3+(t-3)Z_4=1.58$, for $4\leq t<5$, $\widehat{Z}_t=(5-t)Z_4+(t-4)Z_5=0.41$. The graph of $\widehat{Y}_t$ against $t$ is plotted in Figure 1.1.

$\widehat{Y}$ can be seen in a sense as a $G$-Brownian motion, see Ruan [97]. The increments of the $G$-Brownian motion are zero mean, independent and stationary, uncertain variance in the interval $[\underline{\sigma}_t^2,\overline{\sigma}_t^2]$, and can be proved to be $G$-normally distributed $\mathcal{N}(0,[\underline{\sigma}^2,\overline{\sigma}^2])$. We can assume that each experiment about all the urns' composition discussed in the second scenario constitutes a probability measure on $\Omega$. By performing series of experiments, we obtain a set of probability martingale measures $\mathcal{P}_\sigma$ on $\Omega$. Let $\mathcal{P}$ be a

[7]Let $X=(X_1,\ldots,X_n)$ and $Y=(Y_1,\ldots,Y_n)$ be two random variables. They are identically distributed if $\mathcal{E}(\psi(X))=\mathcal{E}(\psi(Y))$ for any $\psi\in C_b(\mathbb{R}^n)$ where $C_b(\mathbb{R}^n)$ is the space of bounded continuous functions.

[8]For any two random variables $X$ and $Y$, $X$ is said to be independent from $Y$ if for each $\psi\in C_b(\mathbb{R}^2)$, $\mathcal{E}(\psi(X,Y))=\mathcal{E}(\mathcal{E}(\psi(x,Y))_{x=X})$.

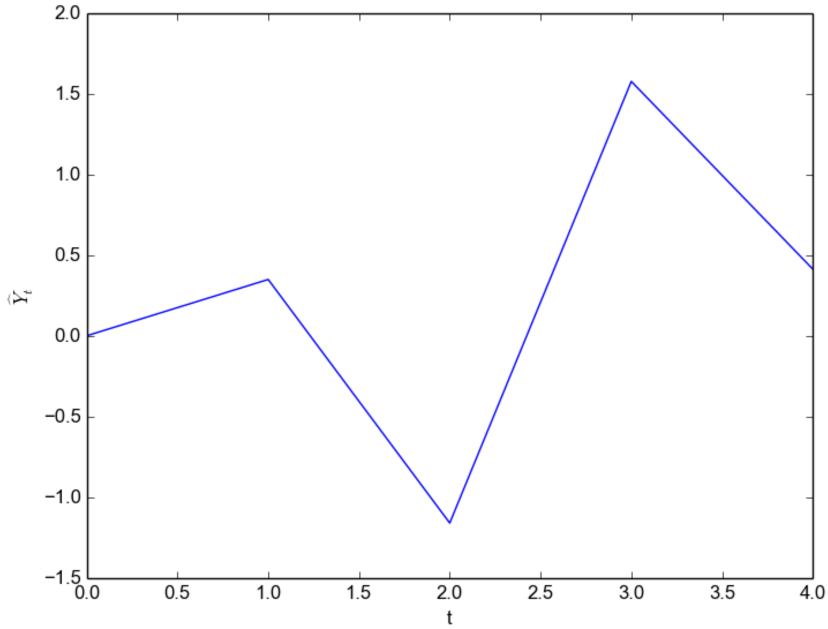


FIGURE 1.1: Graph of $\widehat{Y}_t$ against $t$ .

given set of measures on $\Omega$. For any bounded continuous function $\xi : \Omega \to \mathbb{R}$, if $\sup_{P\in\mathcal{P}_\sigma} \mathbb{E}^P[\xi]$ converges to $\sup_{P\in\mathcal{P}} \mathbb{E}^P[\xi]$, then $\mathcal{P}_\sigma$ converges weakly to $\mathcal{P}$. The supremum of the expectation over $\mathcal{P}$ is the $G$-expectation. In Chapter 2, we discuss this type of convergence in the spirit of Donsker's theorem on a finite state space. We remark that our convergence is much more general than it was suggested by this analogy.

Shige Peng in a series of papers has recently developed a paradigm of probability theory involving sublinear expectation operators. With this new theory, a new type of Itô-integral with respect to the $G$-Brownian motion (see, Peng [87, 89]), Itô formula (see, Li and Peng [69]), martingale representation theorem (see Soner et al. [99], Song [102]), Levy processes (see, Hu and Peng [56]), have emerged. Peng [87] introduced a notion of $G$-expectation and $G$-Brownian motion via a fully one-dimensional nonlinear heat equation[9]. There also exists an alternative representation of the $G$-expectation known as the dual view on $G$-expectation via volatility uncertainty, see Denis et al. [37]. Denis et al. [37] constructed consistent

[9] This is similar to the classical heat equation, dating back to $1900s$, that describes the Brownian motion.

$G$-expectation and $G$-Brownian motion. They also construct the $G$-stochastic integral using quasi-sure stochastic analysis. Cohen et al. [24] presented a theory of sublinear martingales in a discrete setting. By the natural properties of sublinear expectation, one can consider coherent (i.e., sublinear) risk measures[10] as sublinear expectations defined on the space of risk measure, Peng [87]. Thus, the $G$-expectation appears as a natural tool to measure risk under uncertainty, Bion-Nadal and Kervarec [17]. The $G$-Brownian motion provides a powerful tool for modeling path dependent derivatives where volatility uncertainty generates a set of probabilities that are non-equivalent.

Epstein and Ji [43, 44] provided the mathematical foundation for an equilibrium analysis of assets markets with $G$-Brownian stochasticity. They present a model of utility for continuous-time financial markets that captures the agent's concern with $G$-Brownian stochasticity. They also present some no-arbitrage pricing arguments based on hedging strategies. Asset pricing theory characterized by model uncertainty was originally investigated by Avellaneda et al. [7] and Lyons [80]. They characterized the lower and upper bounds of the interval of no-arbitrage prices that generate as the paths of volatility vary in such a confidence interval as a solution to the nonlinear Black-Scholes PDE. This nonlinear equation is associated with the $G$-Brownian motion, see Vorbrink [106]. Avellaneda et al. [7] assert that the presence of volatility uncertainty in a market generates market incompleteness. Thus, perfect hedging is implausible. Accordingly, hedging strategies only yield an interval prediction of asset prices. This bolstered the idea of considering preferences and equilibrium. In the words of Epstein and Ji [43]:

> *"Sharper predictions can be obtained by assuming preference maximization and equilibrium".*

Epstein and Ji [43] analyzed such equilibrium, thus, applied the model of utility to a single agent economy to study equilibrium of asset returns. Denis et al. [37], Cont [25], Vorbrink [106] analyzed hedging strategies to derive a confidence interval of asset prices. Epstein and Ji [43] characterized this asset prices interval in terms of state prices. Beissner [10] proved the existence of Radner equilibrium in an endogenous incomplete market under volatility uncertainty.

[10] Artzner et al. [6] introduced the concept and the axiomatic characterization of sublinear risk measures on finite probability spaces in order to quantify the risk in finance. Delbaen [36] extended Artzner et al. [6] result on general probability spaces.

## 1.2 Bridge between discrete-time and continuous-time models

The connection between a discrete-time stochastic process and a continuous-time stochastic process remains a crucial issue of fundamental importance in stochastic analysis and financial economics. In financial economics, an indepth understanding of the basic economic *arguments* in the discrete-time framework is still the requisite for studying the continuous-time counterpart. However, there has been criticism that the mathematical tools used for analyzing continuous-time financial models have become too complex to capture and simplify the basic economic arguments, see Duffie and Protter [40]. Moreover, financial economists find the discrete-time setting easier to understand in the sense that it captures and gives a simple interpretation to basic economic arguments. Thus, it is important to verify that as one takes the limit of a discrete-time model, when the number of periods increases (goes to infinity), it converges to the continuous-time counterpart. This technique is known as *weak approximation*.

Weak approximation is a very crucial tool in stochastic analysis. A remarkable result is Donsker's theorem (see Chapter 2 for the $G$-Donsker result). For the Donsker-type result for a general class of martingales and diffusion processes, see for example, Billingsley [16]. Many authors have applied the concept of weak approximation to problems in financial economics (see for example, Cox et al. [27], Duffie and Protter [40], Cutland et al. [32] and the references therein), especially after Cox et al. [27] derived the Black-Scholes option pricing formula in an elementary way, as the limit of a discrete-time binomial option pricing formula. This result proves to be more intuitive, elementary and convenient for computation. Despite the wide application of the weak approximation technique to investigate continuous-time models based on discrete analysis, there is no guarantee that such limiting arguments lead to the appropriate continuous-time model, Cutland et al. [30]. Thus, a stronger mode of convergence might be appropriate. The theory of infinitesimals or nonstandard analysis, introduced by Abraham Robinson [94] gives a definitive solution.

### 1.2.1 The theory of infinitesimals

Abraham Robinson [94] developed a rigorous mathematical theory of infinitesimals based on techniques from mathematical logic. His work started with a mathematical object such as the system of real numbers or some Banach space. Robinson and Zakon [95] constructed a nonstandard enlargement even of the full superstructure over the reals. In Chapter 3, we present a more intuitive construction of the nonstandard enlargement. Loeb [76] developed nonstandard measure theory: He showed how every hyperfinite probability space (i.e., a probability space that may be infinite but possesses all the "formal" properties of finite probability spaces) induces a probability space in the standard sense, i.e. a $\sigma$-additive probability measure on some $\sigma$-algebra (viz. the $\sigma$-algebra generated by the internal algebra of the hyperfinite probability space). The corresponding measure on this space is called the *Loeb measure*, see Chapter 3 for an alternative construction of the Loeb measure. Anderson [3] used Loeb's [76] result to develop a hyperfinite construction of the Brownian motion. The hyperfinite Brownian motion can be seen simultaneously as the standard Brownian motion. He also presented the hyperfinite construction of the Brownian stochastic integral. The Itô stochastic integral with respect to the Brownian motion can be constructed as the limit of a pathwise Stieltjes integrals. However, the limit in this construction is not pathwise, but typically $L^2$-limit. The reason for this is that the (standard) Brownian motion has unbounded variation, and the Stieltjes integrals are only defined with respect to paths of bounded variation. Since a hyperfinite random walk is of hyperfinite variation, a Stieltjes integral with respect to hyperfinite random walk can also be seen as a standard stochastic integral in a formal sense, see Anderson [3]. Lindstrøm [74] proposed the notion of a hyperfinite Lévy process and proved that the standard part of a hyperfinite Lévy process is a Lévy process, and that for each infinitesimal generator of a Lévy process one can find a hyperfinite process whose standard part has precisely that generator. This was further discussed in subsequent papers by Albeverio and Herzberg [1] and Ng [84]. Albeverio and Herzberg [1], Hoover and Perkins [54], Herzberg [52], Lindstrøm [70, 71, 72, 75] constructed stochastic integrals with respect to hyperfinite Lévy processes and more general martingales.

In Chapter 4, we present the hyperfinite construction of the $G$-expectation and its corresponding $G$-Brownian motion. We show that our hyperfinite $G$-expectation is infinitely close to the classical $G$-expectation. We remark that we do not work on the Loeb space because the $G$-expectation and its corresponding $G$-Brownian motion is not based on a classical probability measure, but on a set of martingale laws.

The application of nonstandard analysis is not limited to measure theory. Luxemburg [79] developed nonstandard functional analysis using nonstandard hulls [11]. Hurd and Loeb [57] introduced nonstandard analysis to real analysis. An excellent exposition of nonstandard analysis that puts accent on applications in stochastic analysis and mathematical physics is Albeverio et al. [2].

### 1.2.2 Financial markets with nonstandard analysis

The twin frameworks of nonstandard analysis as a continuous-time setting, as well as formally finite setting, motivate its applications to financial economics. Nonstandard measure theory has been successfully applied in studying problems in asset pricing (e.g. Cutland et al. [30], Kopp [64], Cutland et al. [31], Khan and Sun [60]) and in equilibrium theory (e.g. Brown and Robinson [20], Anderson [4], Rashid [92] Anderson and Raimondo [5] and Sun [104].). Anderson's [4] article in the *Handbook of Mathematical Economics* gives a good introduction to the application of nonstandard analysis to economics.

Cutland et al. [30] used nonstandard analysis to construct the Black-Scholes option pricing model. The model can be seen as the hyperfinite version of the binomial Cox-Ross-Rubinstein model [27] and simultaneously as the classical Black-Schole option model. Cutland et al. [31] analyzed the Cox-Ross-Rubinstein jump process option pricing model using nonstandard methods. Kopp [64] used nonstandard methods to establish a link between a mathematical model based on discrete and finite probability spaces and its continuous-time counterpart. They also used nonstandard methods to study the underlying setting. The main concern of these papers is the convergence of discrete-time option pricing models to the continuous-time counterpart.

The application of nonstandard analysis to general equilibrium theory was initiated by the seminal work of Brown and Robinson [20] on nonstandard exchange economies. In their work, they studied equilibria in economies with hyperfinite

[11] This was extended by Henson and Moore [50]. In a sense, the nonstandard hull of an internal Banach space corresponds to the ultraproduct of different Banach spaces.

sets of agents in which each agent has null influence on the economy, and this motivated further applications, e.g. Rashid [92], Anderson and Raimondo [5], Sun [104]. Anderson and Raimondo [5] used nonstandard analysis to prove the existence of equilibrium in a continuous time financial model with Brownian information for multiple agents with the assumption that the dividends are only paid at the expiry date. One way to summarize Anderson and Raimondo's [5] result is as follows: They begin with a standard continuous-time financial model, and then discretize the model to a nonstandard hyperfinite model, they replace the Brownian stochasticity in the continuous-time model with a modified version of Anderson's [3] hyperfinite random walks and introduce the corresponding stochastic integrals. Furthermore, they prove the equilibrium consumption is nonzero at every time and state space. By the construction of the Loeb measure, they produce a candidate for the equilibrium of the aforementioned continuous-time model. Then, they describe the candidate equilibrium as integrals with respect to the normal distribution via the central limit theorem; however, having multiple agents, the dividends depend on the distribution of wealth only at the expiry date. Afterwards, they show that the hyperfinite equilibrium is infinitely close to the candidate equilibrium. Finally, in the spirit of Brown and Robinson [20], they prove that the candidate equilibrium is indeed an equilibrium of the continuous-time model. In the spirit of Anderson and Raimondo [5], Herzberg [53] proved the existence of equilibrium in a continuous-time market with a single agent where the dynamics of the dividends is driven in a sense by hyperfinite Lévy processes.

The theory of hyperfinite $G$-expectation developed in Chapter 4 provides the mathematical foundation that will be needed for the extension of Anderson and Raimondo's [5] result to a continuous-time model driven by $G$-Brownian stochasticity.

# Chapter 2

# Weak Approximation of $G$-Expectation with Discrete State-Space

## 2.1 Introduction

Dolinsky et al. [38] showed a Donsker-type result for the $G$-Brownian motion, henceforth referred to as *G-Donsker,* by introducing a notion of volatility uncertainty in discrete time and defined a discrete version of Peng's $G$-expectation. In the continuous-time limit, the resulting sublinear expectation converges weakly to the $G$-expectation. In their discretization, Dolinsky et al. [38] allow for martingale laws whose support is the whole set of reals. In other words, they only discretized the time line, but not the state space of the canonical process. Now for certain applications, for example a hyperfinite construction of the $G$-expectation in the sense of Robinsonian nonstandard analysis, a discretization of the state space would be necessary. We will show in this chapter that a modification of the construction by Dolinsky et al. [38] suffices to obtain a discretization where the state space for the discrete-time canonical process is discretized, too (whence the martingale laws are supported by a finite lattice only). We will prove the convergence of this discretization to continuous-time $G$-expectation. The proof is based on technique from (linear) probability theory. Ruan [97] constructed the $G$-Brownian motion via the weak limit of a sequence of *G-random walks* which can be seen as the invariance principle of $G$-Brownian motion. The proof relies heavily on the theory of sublinear expectation.

The rest of this chapter is organized as follows: In Section 2.2, we introduce $G$-expectation, the discrete-time and continuous-time version of the sublinear expectation, and the strong formulation of volatility uncertainty in the spirit of Dolinsky et al. [38]. Unlike in Dolinsky et al. [38], we require the discretization of the martingale laws to be defined on a finite lattice rather than the whole set of reals. In Section 2.3, we show that a natural push forward of our discretized sublinear expectation converges weakly to the $G$-expectation as $n$ goes to infinity provided the domain of volatility uncertainty is scaled by $1/n$. Finally, we prove that

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] = \lim_{n\to\infty} \max_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)].$$

## 2.2 Framework

### 2.2.1 $G$-expectation via volatility uncertainty

Peng [89] introduced a sublinear expectation on a well-defined space $\mathbb{L}^1_G$, the completion of $\mathrm{Lip}_{b.cyl}(\Omega)$ (bounded and Lipschitz cylinder function) under the norm $\|\cdot\|_{\mathbb{L}^1_G}$, under which the increments of the canonical process $(B_t)_{t>0}$ are zero-mean, independent and stationary and can be proved to be $(G)$-normally distributed. This type of process is called *$G$-Brownian motion* and the corresponding sublinear expectation is called *$G$-expectation*. We fix a constant $T > 0$ and replace the $d$-dimensional setting by Dolinsky et al. [38] with $d = 1$. We also fix a nonempty, compact and convex set $\mathbf{D} \subseteq \mathbb{R}_+$ such that the volatility processes take values in $\mathbf{D}$.

The $G$-expectation $\xi \mapsto \mathcal{E}^G(\xi)$ is a sublinear operator defined on a class of random variables on $\Omega$. The symbol $G$ refers to a given function

$$G(\gamma) := \frac{1}{2}\sup_{c\in\mathbf{D}} c\gamma : \mathbb{R} \to \mathbb{R} \tag{2.1}$$

where $\mathbf{D} = [r_{\mathbf{D}}, R_{\mathbf{D}}]$ and $0 \le r_{\mathbf{D}} \le R_{\mathbf{D}} < \infty$ are fixed numbers. The construction of the $G$-expectation is as follows. Let $\xi = f(B_T)$, where $B_T$ is the $G$-Brownian motion and $f$ a sufficiently regular function. Then $\mathcal{E}^G(\xi)$ is defined to be the initial value $u(0,0)$ of the solution of the nonlinear backward heat equation,

$$\partial_t u - G(\partial^2_{xx} u) = 0,$$

with terminal condition $u(\cdot, T) = \xi$, Pardoux and Peng [86]. The mapping $\mathcal{E}^G$ can be extended to random variables of the form $\xi = f(B_{t_1}, \cdots, B_{t_n})$ by a stepwise evaluation of the PDE and then to the completion $\mathbb{L}^1_G$ of the space of all such random variables. Denis et al. [37] showed that $\mathbb{L}^1_G$ is the completion of $\mathcal{C}_b(\Omega)$ and $\mathrm{Lip}_{b.cyl}(\Omega)$ under the norm $\|\cdot\|_{\mathbb{L}^1_G}$, and that $\mathbb{L}^1_G$ is the space of the so-called quasi-continuous function and contains all bounded continuous functions on the canonical space $\Omega$, but not all bounded measurable functions are included. Theorem 2.3.13 (our main result in this chapter) cannot be extended to the case where $\xi$ is defined on $\mathbb{L}^1_G$ under the norm $\|\cdot\|_{\mathbb{L}^1_G}$ (see below), thus, we work in a smaller space $\mathbb{L}^1_*$ defined as the completion of $\mathcal{C}_b(\Omega;\mathbb{R})$ under the norm $\|\cdot\|_*$. Our setting is based on a set of martingale laws not a single probability measure. However, when $r_{\mathbf{D}} = R_{\mathbf{D}} = 1$, the canonical process under $\mathcal{E}^G(\cdot)$, $G$-Brownian motion, becomes the standard Brownian motion since $\mathcal{E}^G(\cdot)$ will be a *linear* expectation under the Wiener measure.

There also exists an alternative representation of the $G$-expectation known as the dual view on $G$-expectation via volatility uncertainty, see Denis et al. [37]: One can show that the $G$-expectation can be expressed as the upper expectation

$$\mathcal{E}^G(\xi) = \sup_{P \in \mathcal{P}^G} \mathbb{E}^P[\xi], \quad \xi = f(B_T), \tag{2.2}$$

where $\mathcal{P}^G$ is defined as the set of probability measures on $\Omega$ such that, for any $P \in \mathcal{P}^G$, $B$ is a martingale with the volatility $d\langle B\rangle_t/dt \in \mathbf{D} \quad P \otimes dt$ a.e, and $\mathbf{D} = [r_{\mathbf{D}}, R_{\mathbf{D}}]$, for $0 \leq r_{\mathbf{D}} \leq R_{\mathbf{D}} < \infty$.

**Remark 2.2.1.** (2.2) can be seen as the cheapest super-hedging price of a European contingent claim where $\xi$ can be regarded as the discounted payoff.

### 2.2.2 Continuous-time construction of sublinear expectation

Let $\Omega = \{\omega \in \mathcal{C}([0,T];\mathbb{R}) : \omega_0 = 0\}$ be the canonical space of continuous paths with time horizon $T \in (0,\infty)$, endowed with uniform norm $\|\omega\|_\infty = \sup_{0\leq t\leq T} |\omega_t|$, where the Euclidean norm on $\mathbb{R}$ is given by $|\cdot|$. Let $B$ be the canonical process $B_t(\omega) = \omega_t$, and $\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t)$ is the filtration generated by $B$. A probability measure $P$ on $\Omega$ is called a martingale law provided $B$ is a $P$-martingale and $B_0 = 0$ $P$ a.s. Then, $\mathcal{P}_{\mathbf{D}}$ is the set of martingale laws on $\Omega$ and the volatility takes values in $\mathbf{D}$, $P \otimes dt$ a.e;

$$\mathcal{P}_{\mathbf{D}} = \{P \text{ martingale law on } \Omega\text{: } d\langle B\rangle_t/dt \in \mathbf{D},\ P \otimes dt \text{ a.e.}\}.$$

Thus, the sublinear expectation is given by

$$\mathcal{E}_{\mathbf{D}}(\xi) = \sup_{P \in \mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi], \tag{2.3}$$

such that, for any $\xi : \Omega \to \mathbb{R}$, $\xi$ is $\mathcal{F}_T$-measurable and integrable for all $P \in \mathcal{P}_{\mathbf{D}}$. $\mathbb{E}^P$ denotes the expectation under $P$. It is important to note that the continuous-time sublinear expectation (2.3) can be considered as the $G$-expectation (for every $\xi \in \mathbb{L}^1_G$ where $\mathbb{L}^1_G$ is defined as the $\mathbb{E}[|\cdot|]-$norm completion of $\mathcal{C}_b(\Omega;\mathbb{R})$) provided (2.1) is satisfied (cf. Dolinsky et al. [38]).

### 2.2.3 Discrete-time construction of sublinear expectation

Here we introduce the setting of the discrete-time sublinear expectation. We denote

$$\mathcal{L}_n = \left\{\frac{j}{n\sqrt{n}}, \quad -n^2\sqrt{R_{\mathbf{D}}} \leq j \leq n^2\sqrt{R_{\mathbf{D}}}, \quad \text{for } j \in \mathbb{Z}\right\},$$

and $\mathcal{L}_n^{n+1} = \mathcal{L}_n \times \cdots \times \mathcal{L}_n (n+1 \text{ times})$, for $n \in \mathbb{N}$. Let $X^n = (X^n_k)^n_{k=0}$ be the canonical process $X^n_k(x) = x_k$ defined on $\mathcal{L}_n^{n+1}$ and $(\mathcal{F}^n_k)^n_{k=0} = \sigma(X^n_l, l = 0, \dots, k)$ be the filtration generated by $X^n$. Let

$$\mathbf{D}'_n = \mathbf{D} \cap \left(\frac{1}{n}\mathbb{N}\right)^2$$

be a nonempty bounded set of volatilities. Recall $\mathbf{D} = [r_{\mathbf{D}}, R_{\mathbf{D}}]$, for $0 \leq r_{\mathbf{D}} \leq R_{\mathbf{D}} < \infty$. We note that $R_{\mathbf{D}} = \sup_{\gamma \in \mathbf{D}} |\gamma|$, where $|\cdot|$ denotes the absolute value. A probability measure $P$ on $\mathcal{L}_n^{n+1}$ is called a martingale law provided $X^n$ is a $P$-martingale and $X^n_0 = 0$ $P$ a.s. The increment $\Delta X^n$ denotes the difference by $\Delta X^n_k = X^n_k - X^n_{k-1}$. Let $\mathcal{P}^n_{\mathbf{D}}$ be the set of martingale laws of $X^n$ on $\mathbb{R}^{n+1}$, i.e.,

$$\mathcal{P}^n_{\mathbf{D}} = \left\{P \text{ martingale law on } \mathbb{R}^{n+1}\text{: } r_{\mathbf{D}} \leq |\Delta X^n_k|^2 \leq R_{\mathbf{D}},\ P \text{ a.s.}\right\},$$

such that for all $n$, $\mathcal{L}_n^{n+1} \subseteq \mathbb{R}^{n+1}$.

In order to establish a relation between the continuous-time and discrete-time settings, we obtained a continuous-time process $\widehat{x}_t \in \Omega$ from any discrete path $x \in \mathcal{L}_n^{n+1}$ by linear interpolation. i.e.,

$$\widehat{x}_t := (\lfloor nt/T \rfloor + 1 - nt/T) x_{\lfloor nt/T \rfloor} + (nt/T - \lfloor nt/T \rfloor) x_{\lfloor nt/T \rfloor + 1}$$

where $\widehat{\ }: \mathcal{L}_n^{n+1} \to \Omega$ is the linear interpolation operator, $x = (x_0, \ldots, x_n) \mapsto \widehat{x} = \{(\widehat{x})_{0 \le t \le T}\}$, and $\lfloor y \rfloor$ denotes the greatest integer less than or equal to $y$. If $X^n$ is the canonical process on $\mathcal{L}_n^{n+1}$ and $\xi$ is a random variable on $\Omega$, then $\xi(\widehat{X}^n)$ defines a random variable on $\mathcal{L}_n^{n+1}$.

**Remark 2.2.2.** If $n = T$, thus for all $t \in \mathbb{N}$,

$$\widehat{x}_t := (\lfloor t \rfloor + 1 - t) x_{\lfloor t \rfloor} + (t - \lfloor t \rfloor) x_{\lfloor t \rfloor + 1}.$$

For instance, if $x_0 = 0$, $x_1 = 2$, $x_2 = 1$, $x_3 = 5$, $x_4 = 3$, for $0 \le t < 1$, $\widehat{x}_t = (1 - t)x_0 + tx_1 = 0$; for $1 \le t < 2$, $\widehat{x}_t = (2 - t)x_1 + (t - 1)x_2 = 2$; for $2 \le t < 3$, $\widehat{x}_t = (3 - t)x_2 + (t - 2)x_3 = 1$; for $3 \le t < 4$, $\widehat{x}_t = (4 - t)x_3 + (t - 3)x_4 = 5$, for $4 \le t < 5$, $\widehat{x}_t = (5 - t)x_4 + (t - 4)x_5 = 3$. Thus, the graph of $\widehat{x}_t$ as a function of $t$ is plotted in Figure 2.1.

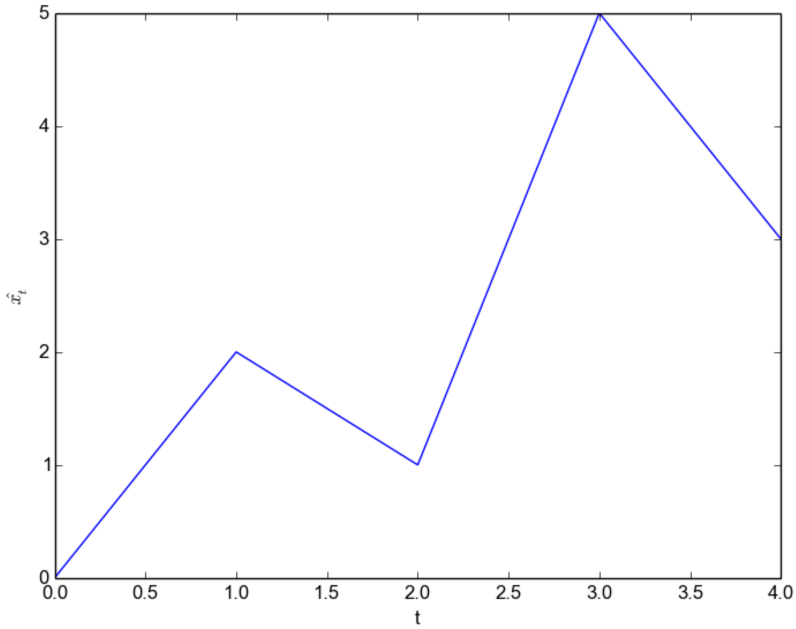


FIGURE 2.1: Graph of $\widehat{x}_t$ against $t$

### 2.2.4 Strong formulation of volatility uncertainty

We introduce the so-called strong formulation of volatility uncertainty for the continuous-time construction, as in Dolinsky et al. [38], Nutz [85], Soner et al. [100, 101], and for the discrete-time construction, as in Dolinsky et al. [38]; i.e., we consider martingale laws generated by stochastic integrals with respect to a fixed Brownian motion and a fixed random walk.

For the continuous-time construction; let $\mathcal{Q}_{\mathbf{D}}$ be the set of martingale laws of the form:

$$\mathcal{Q}_{\mathbf{D}} = \left\{ P_0 \circ (M)^{-1};\ M = \int f(t, B) dB_t, \text{ and } f \in \mathcal{C}([0,T] \times \Omega; \sqrt{\mathbf{D}}) \text{ is adapted} \right\}.$$

$B$ is the canonical process under the Wiener measure $P_0$, and $\mathbf{D}$ is a convex set.

**Remark 2.2.3.** The elements of $\mathcal{Q}_{\mathbf{D}}$, in particular $M$, with nondegenerate $f$ which satisfies the predictable representation condition, correspond to the analogy of market completeness in finance (martingale representation theorem).

For the discrete-time construction; we fix $n \in \mathbb{N}$, $\Omega_n = \{\omega = (\omega_1, \dots, \omega_n) : \omega_i \in \{\pm 1\}, \quad i = 1, \dots, n\}$ equipped with the power set and let

$$P_n = \underbrace{\frac{\delta_{-1} + \delta_{+1}}{2} \otimes \cdots \otimes \frac{\delta_{-1} + \delta_{+1}}{2}}_{\text{n times}}$$

where for all $A \subseteq \mathbb{R}$,

$$\delta_x(A) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

be the product probability associated with the uniform distribution. Let $\xi_1, \dots, \xi_n$ be an i.i.d sequence of $\{\pm 1\}$ -valued random variables. The components of $\xi_k$ are orthonormal in $L^2(P_n)$. We denote the associated random walk by

$$Z_k^n = \sum_{l=1}^{k} \xi_l,$$

then, we can view

$$\sum_{l=1}^{k} f(l-1, \mathbb{X}) \Delta \mathbb{X}_l$$

as the discrete-time stochastic integrals of $\mathbb{X}$, where $f$ is $\mathcal{F}^n$-adapted and

$$\mathbb{X} = \frac{1}{\sqrt{n}} Z^n$$

is the scaled random walk. We denote by $\mathcal{Q}^n_{\mathbf{D}'_n}$ the set of martingale laws of the form:

$$\mathcal{Q}^n_{\mathbf{D}'_n} = \left\{ P_n \circ (M^{f,\mathbb{X}})^{-1};\ f : \{0, \dots, n\} \times \mathcal{L}_n^{n+1} \to \sqrt{\mathbf{D}'_n} \text{ is } \mathcal{F}^n\text{-adapted.} \right\} \quad (2.4)$$

where

$$M^{f,\mathbb{X}} = \left( \sum_{l=1}^{k} f(l-1, \mathbb{X}) \Delta \mathbb{X}_l \right)_{k=0}^{n}.$$

## 2.3 Results and proofs

Theorem 2.3.1 states that a sublinear expectation with discrete-time volatility uncertainty on a set of reals converges to the $G$-expectation.

**Theorem 2.3.1.** (cf. Dolinsky et al. [38, Theorem 2.2]) Let $\xi : \Omega \to \mathbb{R}$ be a continuous function satisfying $|\xi(\omega)| \leq a(1+ \| \omega \|_\infty)^b$ for some constants $a, b > 0$. Then,

$$\lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{P}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] = \sup_{P\in\mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi]. \tag{2.5}$$

To prove (2.5), we prove two separate inequalities which imply (2.5). The proof of the first inequality (for $\leq$ in (2.5)) will be discussed below while the proof of the second inequality (for $\geq$ in (2.5)) is embedded in the proof of Proposition 2.3.11. Before then, we introduce a smaller space $\mathbb{L}^1_*$ that is defined as the completion of $\mathcal{C}_b(\Omega;\mathbb{R})$ under the norm

$$\| \xi \|_* := \sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[|\xi|], \quad \mathcal{Q} := \mathcal{P}_{\mathbf{D}} \cup \{P\circ(\widehat{X}^n)^{-1}; P\in\mathcal{P}^n_{\mathbf{D}/n}, n\in\mathbb{N}.\}.$$

This is because Theorem 2.3.1 will not hold if $\xi$ just belong to $\mathbb{L}^1_G$, where $\mathbb{L}^1_G$ is the completion of $\mathcal{C}_b(\Omega;\mathbb{R})$ under the norm

$$\| \xi \|_G := \sup_{P\in\mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[|\xi|]. \tag{2.6}$$

In fact, a random variable which is defined on a set of paths of finite variation will have zero expectation under any martingale law $P \in \mathcal{P}_{\mathbf{D}}$ because the support of the martingale laws is disjoint to a set of paths of finite variation whereas it will have non zero expectation under an element of $\mathcal{Q}$.

**Lemma 2.3.2.** (cf. Dolinsky et al. [38, Lemma 3.4]) Let $\xi : \Omega \to \mathbb{R}$ be a continuous function satisfying $|\xi(\omega)| \leq a(1+ \| \omega \|_\infty)^b$ for some constants $a, b > 0$. Then, $\xi \in \mathbb{L}^1_*$.

We shall prove Lemma 2.3.2 later.

To prove the first inequality (for $\leq$ in (2.5)),

$$\limsup_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{P}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \leq \sup_{P\in\mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi], \tag{2.7}$$

we need to prove the following lemmas. The main idea is to establish the stability result for the quadratic variation of the process $B$. This will be done in Lemma 2.3.6, and the necessary tightness condition is as a result of the compactness of $\mathbf{D}$ in the sense that $R_{\mathbf{D}}$ is finite.

**Lemma 2.3.3.** (cf. Dolinsky et al. [38, Lemma 3($i$) and ($ii$)]) For any given $1 \le q < \infty$, there exists a positive constant $C$ such that for all $0 \le k \le l \le n$ and $\mathbb{Q} \in \mathcal{P}^n_{\mathbf{D}}$,

$$\mathbb{E}^{\mathbb{Q}}[\sup_{k=0,\dots,n} |X^n_k|^{2q}] \le C(nR_{\mathbf{D}})^q \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}[|X^n_l - X^n_k|^4] \le CR^2_{\mathbf{D}}(l-k)^2.$$

*Proof.* The main idea is the Burkholder-Davis-Gundy inequalities: For any given $1 \le q < \infty$, and a positive universal constant $K = K(q,1)$,

$$\mathbb{E}^{\mathbb{Q}}[\sup_{k=0,\dots,n} |X^n_k|^{2q}] \le K\mathbb{E}^{\mathbb{Q}}[|\langle X\rangle_n|^q].$$

Since $\mathbb{Q} \in \mathcal{P}^n_{\mathbf{D}}$, $|\langle X\rangle_n| = |\sum_{l=1}^n (\Delta X^n_l)^2| \le nR_{\mathbf{D}}$ $\mathbb{Q}$ a.s.
With similar argument,

$$\mathbb{E}^{\mathbb{Q}}[|X^n_l - X^n_k|^4] \le K\mathbb{E}^{\mathbb{Q}}[|\langle X\rangle_l - \langle X\rangle_k|^2],$$

then, $\mathbb{Q} \in \mathcal{P}^n_{\mathbf{D}}$ implies that $|\langle X\rangle_l - \langle X\rangle_k|^2 \le (l-k)^2 R^2_{\mathbf{D}}$ $\mathbb{Q}$ a.s.
It is important to note that if $\gamma_1, \gamma_2, \dots, \gamma_m \in \mathbf{D}$, then $\gamma_1 + \gamma_2 + \ldots + \gamma_m \in m\mathbf{D}$ by the convexity of $\mathbf{D}$. $\square$

The next lemma shows that all the expressions in (2.5) are well-defined and finite.

**Lemma 2.3.4.** (cf. Dolinsky et al. [38, Lemma 3.2]) Let $\xi : \Omega \to \mathbb{R}$ be a continuous function satisfying $|\xi(\omega)| \le a(1+ \parallel \omega \parallel_\infty)^b$ for some constants $a, b > 0$. Then, $\parallel \xi \parallel_*$ is finite. i.e.,

$$\sup_{n\in\mathbb{N}} \sup_{\mathbb{Q}\in\mathcal{P}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[|\xi(\widehat{X}^n)|] < \infty \tag{2.8}$$

and

$$\sup_{P\in\mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[|\xi|] < \infty. \tag{2.9}$$

*Proof.* Using the condition on $\xi$, for any $1 \le q < \infty$, $n \in \mathbb{N}$ and $\mathbb{Q} \in \mathcal{P}^n_{\mathbf{D}/n}$,

$$\mathbb{E}^{\mathbb{Q}}[|\xi(\widehat{X}^n)|] \le a + a\mathbb{E}^{\mathbb{Q}}\left[\sup_{0\le t\le T} |\widehat{X}^n_t|^q\right] \le a + a\mathbb{E}^{\mathbb{Q}}\left[\sup_{k=0,\dots,n} |X^n_k|^q\right].$$

By Lemma 2.3.3, $\mathbb{E}^{\mathbb{Q}}[\sup_{k=0,\dots,n} |X^n_k|^{2q}] \le K(nR_{\mathbf{D}})^q$ and as we observe that $R_{\mathbf{D}/n} = R_{\mathbf{D}}/n$,

$$\mathbb{E}^{\mathbb{Q}}[|\xi(\widehat{X}^n)|] \le CR^{q/2}_{\mathbf{D}}.$$

Thus, (2.8) follows. (2.9) also follows from the Burkholder-Davis-Gundy inequalities and from the condition on $\mathbf{D}$ ($\mathbf{D}$ is bounded), i.e.,

$$\mathbb{E}^P[\sup_{0\le t\le T}|B_t|^q] \le K_q \quad \text{for all } P \in \mathcal{P}_{\mathbf{D}}.$$

□

Now, we want to prove the tightness result.

**Lemma 2.3.5.** (cf. Dolinsky et al. [38, Lemma 3.3($i$)]) Let $M^n = (M^n_k)^n_{k=0}$ be a martingale with law $\widetilde{\mathbb{Q}}^n$ in $\mathcal{P}^n_{\mathbf{D}/n}$ on $\mathbb{R}^{n+1}$ and let $\widehat{M}^n$ be martingale with law $Q^n$ on $\Omega$, for each $n \in \mathbb{N}$. Then, the sequence $(Q^n)_n$ is tight on $\Omega$.

*Proof.* Let $0 \le s \le t \le T$. Using that $R_{\mathbf{D}/n} = R_{\mathbf{D}}/n$ and from Lemma 2.3.3,

$$\mathbb{E}^{Q^n}[|B_t - B_s|^4] = \mathbb{E}^{\widetilde{\mathbb{Q}}^n}[|\widehat{M}^n_t - \widehat{M}^n_s|^4] \le K(t-s)^2$$

for a positive constant $K$. By the Kolmogorov's criterion moment for weak relative compactness, (see Klenke [61, Theorem 21.42]), $(Q^n)_n$ is tight. □

**Lemma 2.3.6.** (cf. Dolinsky et al. [38, Lemma 3.3($ii$)]) Let $M^n = (M^n_k)^n_{k=0}$ be a martingale with law $\widetilde{\mathbb{Q}}^n$ in $\mathcal{P}^n_{\mathbf{D}/n}$ on $\mathbb{R}^{n+1}$ and let $\widehat{M}^n$ be martingale with law $Q^n$ on $\Omega$, for each $n \in \mathbb{N}$. Then, any cluster point $Q$ of the sequence $(Q^n)_n$ is an element in $\mathcal{P}_{\mathbf{D}}$.

*Proof.* Let $Q$ be a cluster point. $B$ (in Lemma 2.3.5) is $Q$-martingale as a result of the uniform integrability condition by Lemma 2.3.3. Now, it remains to prove that $d\langle B\rangle_t/dt \in \mathbf{D}$ holds $Q \otimes dt$ a.e. The main technique used here is the separating hyperplane theorem which implies that for $\gamma \in \mathbb{R}_{>0}$,

$$\gamma \in \mathbf{D} \iff f(\gamma) \le \sup_{\beta\in\mathbf{D}} f(\beta) =: K^f_{\mathbf{D}} \tag{2.10}$$

where $f : \mathbb{R} \to \mathbb{R}$ is an arbitrary linear function. Let $\Delta Y_t = Y_t - Y_s$ for fix $0 \le s \le t \le T$. Consider an arbitrary continuous and adapted function $U : [0,T] \times \Omega \to [0,1]$. We then introduce an arbitrary neighborhood $\widetilde{\mathbf{D}}$ of $\mathbf{D}$. Then we have for sufficiently large $n$,

$$\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[(\Delta\widehat{M}^n_t)^2 \mid \sigma(\widehat{M}^n_u : 0 \le u \le s-\varepsilon)] \in (t-s)\widetilde{\mathbf{D}} \quad \widetilde{\mathbb{Q}}^n\text{a.s.}$$

It follows from (2.10) that

$$\begin{aligned}&\mathbb{E}^{Q^n}[U(s-\varepsilon,B)\{f((\Delta B_t)^2)-K^f_{\mathbf{D}}(t-s)\}]\\=&\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[U(s-\varepsilon,\widehat{M}^n)\{f((\Delta\widehat{M}^n_t)^2)-K^f_{\mathbf{D}}(t-s)\}].\end{aligned}$$

Then,

$$\begin{aligned}&\mathbb{E}^{Q^n}[U(s-\varepsilon,B)\{f((\Delta B_t)^2)-K^f_{\mathbf{D}}(t-s)\}]\\=&\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[U(s-\varepsilon,\widehat{M}^n)\{f((\Delta\widehat{M}^n_t)^2)-K^f_{\mathbf{D}}(t-s)\}]\\=&\mathbb{E}[\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[U(s-\varepsilon,\widehat{M}^n)(f((\Delta\widehat{M}^n_t)^2)-K^f_{\mathbf{D}}(t-s))]\mid\sigma(\widehat{M}^n_u:0\le u\le s-\varepsilon)]\\=&\mathbb{E}[U(s-\varepsilon,\widehat{M}^n)\underbrace{\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[(f((\Delta\widehat{M}^n_t)^2)-K^f_{\mathbf{D}}(t-s))]\mid\sigma(\widehat{M}^n_u:0\le u\le s-\varepsilon)]}_{=\underbrace{\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[f((\Delta\widehat{M}^n_t)^2)\mid\sigma(\widehat{M}^n_u:0\le u\le s-\varepsilon)]-K^f_{\mathbf{D}}(t-s)}_{\overset{t-s}{=}\underbrace{f(\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[|\Delta\widehat{M}^n_t|^2\mid\sigma(\widehat{M}^n_u:0\le u\le s-\varepsilon)])-K^f_{\mathbf{D}}(t-s)}_{\overset{t-s}{\le}(\sup_{\widetilde{\mathbf{D}}}f-K^f_{\mathbf{D}})(t-s).}}}]\end{aligned}$$

Since $\widetilde{\mathbf{D}}$ is an arbitrary neighborhood of $\mathbf{D}$, $\sup_{\beta\in\widetilde{\mathbf{D}}}f(\beta)$ will be arbitrary close to $\sup_{\beta\in\mathbf{D}}f(\beta)=K^f_{\mathbf{D}}$, thus,

$$\sup_{\beta\in\widetilde{\mathbf{D}}}f(\beta)-K^f_{\mathbf{D}}$$

is arbitrary close to zero. Hence,

$$\begin{aligned}&\limsup_{n\to\infty}\mathbb{E}^{Q^n}[U(s-\varepsilon,B)\{f((\Delta B_t)^2)-K^f_{\mathbf{D}}(t-s)\}]\\=&\limsup_{n\to\infty}\mathbb{E}^{\widetilde{\mathbb{Q}}^n}[U(s-\varepsilon,\widehat{M}^n)\{f((\Delta\widehat{M}^n_t)^2)-K^f_{\mathbf{D}}(t-s)\}]\le 0.\end{aligned}$$

By the integrability condition of Lemma 2.3.3, these sequences of means converge, whence the lim sup is actually a limit and in each case said to be

$$\mathbb{E}^Q[U(s-\varepsilon,B)f((\Delta B_t)^2)]\le\mathbb{E}^Q[U(s-\varepsilon,B)K^f_{\mathbf{D}}(t-s)].\tag{2.11}$$

Since $U(s-\varepsilon,B)$ is $\mathcal{F}_s$-measurable and $\mathbb{E}^Q[(\Delta B_t)^2)\mid\mathcal{F}_s]=\mathbb{E}^Q[\langle B\rangle_t-\langle B\rangle_s|\mathcal{F}_s]$ (note that $B$ is square martingale, $Q$-martingale), (2.11) equals

$$\mathbb{E}^Q[U(s-\varepsilon,B)f(\langle B\rangle_t-\langle B\rangle_s)]\le\mathbb{E}^Q[U(s-\varepsilon,B)K^f_{\mathbf{D}}(t-s)].\tag{2.12}$$

By the continuity condition of $U$ and the dominated convergence theorem as $\varepsilon \to 0$, (2.12) becomes

$$\mathbb{E}^Q[U(s,B)f(\langle B\rangle_t - \langle B\rangle_s)] \le \mathbb{E}^Q[U(s,B)K^f_{\mathbf{D}}(t-s)]$$

and then

$$\mathbb{E}^Q\left[\int_0^T U(s,B)f(d\langle B\rangle_t)\right] \le \mathbb{E}^Q\left[\int_0^T U(s,B)K^f_{\mathbf{D}}dt\right].$$

By Lemma 2.3.7 (see below for $\Omega' = [0,T]\times\Omega$ and $\mu = Q\otimes dt$), it follows that

$$f(d\langle B\rangle_t/dt) \le K^f_{\mathbf{D}} \quad \text{holds } Q\otimes dt \text{ a.e.}$$

Since $f$ is arbitrary, (2.10) implies that

$$\frac{d\langle B\rangle_t}{dt} \in \mathbf{D} \quad \text{holds } Q\otimes dt \text{ a.e.}$$

$\square$

**Lemma 2.3.7.** If there exists some $C \in \mathbb{R}_{>0}$ and a measure $\mu$ on $\Omega'$ such that for all measurable $f : \Omega' \to [0,1]$,

$$\int Xfd\mu \le C\int fd\mu,$$

then $X \le C$ $\mu$ a.s.

*Proof.* We prove by contradiction. Let assume that there exists some measurable $A \subseteq \Omega'$ such that $\mu(A) > 0$ and $X > C$ on $A$. Put $f = \mathbf{1}_A$. Then,

$$\int X\mathbf{1}_A d\mu \ge \int_A Xd\mu > C\mu(A) = C\int \mathbf{1}_A d\mu,$$

which contradicts the claim. $\square$

We can now prove Lemma 2.3.2.

*Proof of Lemma 2.3.2.* Let $\xi^m = (\xi \wedge m) \vee m$. We want to show that the upper expectation

$$\mathcal{E}_{\mathbf{D}}(\cdot) = \sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[\cdot]$$

is continuous along the decreasing sequence $|\xi - \xi^m|$ where

$$\mathcal{Q} := \mathcal{P}_{\mathbf{D}} \cup \{P \circ (\widehat{X}^n)^{-1}; P \in \mathcal{P}^n_{\mathbf{D}/n}, n \in \mathbb{N}.\}.$$

By the proof of Lemma 2.3.5 we can say that $\mathcal{Q}$ is tight and by the proof of Lemma 2.3.4, $\parallel \xi \parallel_* < \infty$. i.e.,

$$\mathcal{E}_{\mathbf{D}}(\cdot) = \sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[\cdot] < \infty.$$

For a decreasing sequence $|\xi - \xi^m|$,

$$\sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[|\xi - \xi^m|] = \sup_{Q \in \mathcal{Q}} \int_0^\infty Q(|\xi - \xi^m| \geq t) dt.$$

For each fixed $t > 0$, the set $\{|\xi - \xi^m| \geq t\}$ is closed and $\{|\xi - \xi^m| \geq t\} \downarrow \emptyset$ as $m \to \infty$. By Denis et al. [37, Lemma 7]: $\mathcal{Q}$ is relatively compact if and only if for each sequence $Y_n \downarrow \emptyset$, we have $\sup_{Q \in \mathcal{Q}} Q(Y_n) \downarrow 0$. Thus,

$$\sup_{Q \in \mathcal{Q}} \mathbb{E}^Q[|\xi - \xi^m|] \downarrow 0.$$

□

Finally, we can now prove the first inequality of Theorem 2.3.1.

*Proof of Theorem 2.3.1 (for $\leq$ in* (2.5)*).* Let $\xi : \Omega \to \mathbb{R}$ be a continuous function satisfying $|\xi(\omega)| \leq a(1 + \parallel \omega \parallel_\infty)^b$ for some constants $a, b > 0$ and let $\varepsilon > 0$. Let there exists an $\varepsilon$-optimizer $\mathbb{Q}^n \in \mathcal{P}^n_{\mathbf{D}/n}$, that is, if $Q^n$ is the martingale law of $\widehat{X}^n$ on $\Omega$ under $\mathbb{Q}^n$ for each $n \in \mathbb{N}$, then

$$\mathbb{E}^{Q^n}[\xi] = \mathbb{E}^{\mathbb{Q}^n}[\xi(\widehat{X}^n)] \geq \sup_{\mathbb{Q} \in \mathcal{P}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] - \varepsilon.$$

By Lemma 2.3.5 and Lemma 2.3.6, the sequence $(Q^n)_n$ is tight and any cluster point of $(Q^n)_n$ is an element of $\mathcal{P}_{\mathbf{D}}$. Since $\xi$ is continuous and Lemma 2.3.4 implies that $\sup_n \mathbb{E}^{Q^n}[\xi]$ is finite, tightness yields that

$$\sup_{P \in \mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi] \geq \limsup_{n \to \infty} \mathbb{E}^{Q^n}[\xi].$$

Thus,

$$\sup_{P \in \mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi] + \varepsilon \geq \limsup_{n \to \infty} \sup_{\mathbb{Q} \in \mathcal{P}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)].$$

For arbitrary $\varepsilon > 0$,

$$\limsup_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{P}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \le \sup_{P\in\mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi].$$

□

**Proposition 2.3.8.** (cf. Dolinsky et al. [38, Proposition 3.5]) The convex hull of $\mathcal{Q}_{\mathbf{D}}$ is a weakly dense subset of $\mathcal{P}_{\mathbf{D}}$.

*Proof.* This proof is divided into three parts, discretization, randomization technique and the smoothing part.

- Discretization: First, let us recall that

$$\mathcal{P}_{\mathbf{D}} = \left\{P \text{ martingale law on } \Omega\text{: } d\left\langle B\right\rangle_t/dt \in \mathbf{D},\ P\otimes dt \text{ a.e.}\right\}.$$

Let $\bar{M}$ be a martingale whose law lies in $\mathcal{P}_{\mathbf{D}}$ such that

$$\bar{M} = \int \alpha_t dW_t \quad \text{with} \quad \alpha := \sqrt{d\langle\bar{M}\rangle_t/dt} \quad \text{and} \quad W := \int \alpha_t^{-1} d\bar{M}_t.$$

To see that $W$ is a Brownian motion, it suffices to show that $W$ is a local martingale such that $\langle W\rangle_t = t$. From the definition of $W$, $W$ is a local martingale since $\bar{M}$ is a martingale, and $W_0 = 0$. Then,

$$\langle W\rangle_t = \int_0^t \frac{ds}{d\langle\bar{M}\rangle_s} d\langle\bar{M}\rangle_s = \int_0^t ds = t.$$

Thus, $\langle W\rangle_t = t$, and then $W$ is a Brownian motion. Fix $n\ge 1$, we consider

$$\bar{M}^{(n)} = \int \alpha_t^{(n)} dW_t,$$

where $\alpha^{(n)}$ is real-valued piecewise constant process that satisfies

$$(\alpha_t^{(n)})^2 = \lambda_{\mathbf{D}}(y_t^2),$$

for

$$y_t = \frac{n}{T}\int_{(k-1)T/n}^{kT/n} \alpha_s ds, \quad t\in(kT/n,(k+1)T/n],$$

$k = 1, \dots, n-1$ and $\lambda_{\mathbf{D}} : \mathbb{R} \to \mathbf{D}$ the Euclidean projection. We can set, for example, $\alpha^{(n)} = \sqrt{\gamma}$, for $\gamma \in \mathbf{D}$ on $[0, T/n]$, then we have

$$\lim_{n\to\infty} \mathbb{E}^P[|\langle \bar{M}^{(n)} - \bar{M}\rangle_T|] = 0, \quad \lim_{n\to\infty} \mathbb{E}^P\left[\int_0^T |\alpha_t^{(n)} - \alpha_t|^2 dt\right] = 0,$$

and thus $\bar{M}^{(n)}$ converges weakly to $\bar{M}$. Next, we want to show that for a given measure, the measure is in the convex hull of measures of Brownian martingales.

- Randomization technique: We fix a filtered probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{F}}, \bar{\mathbb{P}})$ where $\bar{\mathbb{F}} = (\bar{\mathcal{F}}_t)_{0\le t\le T}$. Let $\bar{M}$ be a $\bar{\mathbb{F}}$-martingale whose law belongs to $\mathcal{P}_{\mathbf{D}}$ and is of the form:

$$\bar{M} = \int \alpha_t dW_t,$$

where

$$\alpha_t = \sum_{k=0}^{n-1} \mathbf{1}_{[t_k, t_{k+1})} \alpha(k),$$

for some time discretization $0 = t_0 < \cdots < t_n = T$ and $W$ is a Brownian motion on the filtered probability space. We consider another probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ equipped with a Brownian motion $\tilde{W}$ and a sequence of i.i.d uniformly distributed random variables $(\tilde{V}_k)_{1\le k\le n}$ which is independent of $\tilde{W}$. By the existence of regular conditional probability measures, we can construct random variables $\tilde{\alpha}(k)$ which are

$$\sigma(\tilde{W}_s, 0 \le s \le t_k) \vee \sigma(\tilde{V}_j, 1 \le j \le k) - \text{ measurable}$$

and such that the law of $(\tilde{W}_s, (\tilde{\alpha}(i))_{0\le i\le n-1})$ under $\tilde{\mathbb{P}}$ equals the law of $(W_s, (\alpha(i))_{0\le i\le n-1})$ under $\bar{\mathbb{P}}$. We can now consider the volatility corresponding to a fixed realization of $\tilde{V}_1, \cdots, \tilde{V}_n$. Indeed, for $v = (v_1, \cdots, v_n) \in (0,1)^n$, let $\tilde{\alpha}(k; v)$ be a random variable which is

$$\sigma(\tilde{W}_s, 0 \le s \le t_k) \vee \sigma(v_j, 1 \le j \le k) - \text{ measurable} \tag{2.13}$$

and consider

$$\tilde{M}^v := \int \left(\sum_{k=0}^{n-1} \mathbf{1}_{[t_k, t_{k+1})} \tilde{\alpha}(k; v)\right) d\tilde{W}_t.$$

Further, we denote $\tilde{P}_v = \tilde{\mathbb{P}} \circ (\tilde{M}^v)^{-1}$. Consider a family of conditional probability measures $(\mathbb{P}_v)_{v\in(0,1)^n}$ of $\tilde{\mathbb{P}}$ with respect to $\sigma(\tilde{V}_k, 1 \le k \le n)$ and define

$\tilde{\mathbb{P}}_v = \mathbb{P}_v \circ (\tilde{M}^v)^{-1}$. For a given bounded continuous function $F$,

$$\begin{aligned}\mathbb{E}^{\bar{\mathbb{P}}}[F(\bar{M})] &= \mathbb{E}^{\tilde{P}_v}[F(\tilde{M}^{\tilde{V}_1,\cdots,\tilde{V}_n})] \\ &= \int_{(0,1)^n} \mathbb{E}^{\tilde{P}_v}[F(\tilde{M}^v)]dv \leq \sup_{v\in(0,1)^n} \mathbb{E}^{\tilde{\mathbb{P}}_v}[F(\tilde{M}^v)].\end{aligned}$$

Thus, by the Hahn-Banach theorem, the law of $\bar{M}$ is contained in the weak closure of the convex hull of the laws of $S$, where $S = \{\tilde{M}^v : v \in (0,1)^n\}$. For each fixed $v$, we note that $\tilde{M}^v = \int h(t,\tilde{W})d\tilde{W}_t$ where $h$ is a measurable, adapted, $\sqrt{\mathbf{D}}$-valued function. Next, we want to show that these Brownian martingales can be approximated by Brownian martingales with some regularity conditions.

- Smoothing: Recall

$$\mathcal{Q}_{\mathbf{D}} = \left\{P_0 \circ (M)^{-1};\ M = \int f(t,B)dB_t,\ \text{and } f \in \mathcal{C}([0,T]\times\Omega;\sqrt{\mathbf{D}}) \text{ is adapted}\right\}.$$

Now, we want to approximate $h$ by a continuous function $f$. Let $h : [0,T]\times\Omega \to \sqrt{\mathbf{D}}$ be a measurable adapted function and $\varepsilon > 0$. Let $\tilde{f} \in C([0,T]\times\Omega;\mathbb{R})$ such that

$$\mathbb{E}\left[\int_0^T |\tilde{f}(t,\tilde{W}) - h(t,\tilde{W})|^2 dt\right] \leq \varepsilon.$$

Such $\tilde{f}$ exists, due to standard density arguments.
Let $f(t,x) := \sqrt{\lambda_{\mathbf{D}}(\tilde{f}(t,x)^2)}$. Then $f \in C([0,T]\times\Omega;\sqrt{\mathbf{D}})$ and

$$|f-h|^2 \leq |f^2-h^2| \leq |\tilde{f}^2-h^2| \leq (|\tilde{f}|+|h|)|\tilde{f}-h| \leq 2\sqrt{R_{\mathbf{D}}}|\tilde{f}-h|.$$

Using Jensen's inequality, we can say that

$$\mathbb{E}\left[\int_0^T |f(t,\tilde{W}) - h(t,\tilde{W})|^2 dt\right] \leq 2\sqrt{TR_{\mathbf{D}}\varepsilon},$$

which, as a result of the above steps, completes the proof.

□

**Remark 2.3.9.** (cf. Dolinsky et al. [38, Remark 3.6 ]) For any bounded continuous $\xi : \Omega \to \mathbb{R}$,

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] = \sup_{P\in\mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi].$$

This follows directly from Proposition 2.3.8. Denis et al. [37], Section 3, show that the $G$-expectation as noted in Peng [87, 88] coincides with the mapping $\xi \mapsto \sup_{P\in\mathbf{Q_D}} \mathbb{E}^P[\xi]$ for a given set $\mathbf{Q_D}$ which satisfies $\mathcal{Q}_\mathbf{D} \subseteq \mathbf{Q_D} \subseteq \mathcal{P}_\mathbf{D}$.

**Lemma 2.3.10.** Let

$$\mathcal{Q}^n_\mathbf{D} = \left\{ P_n \circ \left(M^{f,\mathbb{X}}\right)^{-1};\ f : \{0,\dots,n\} \times \mathbb{R}^{n+1} \to \sqrt{\mathbf{D}} \text{ is adapted.} \right\}$$

where

$$M^{f,\mathbb{X}} = \left( \sum_{l=1}^{k} f(l-1,\mathbb{X}) \Delta \mathbb{X}_l \right)_{k=0}^{n}.$$

Then $\mathcal{Q}^n_\mathbf{D} \subseteq \mathcal{P}^n_\mathbf{D}$.

*Proof.* From the above equation, we can say that $\Delta M^f_k = f(k,\mathbb{X})\xi_k$. And by the orthonormality property of $\xi_k$, we have

$$\mathbb{E}^{P_n}[f(k,\mathbb{X})^2 \xi_k^2 | \mathcal{F}^n_k] = \mathbb{E}^{P_n}[f(k,\mathbb{X})^2 | \mathcal{F}^n_k] \leq \mathbb{E}^{P_n}[(\sqrt{R_\mathbf{D}})^2 | \mathcal{F}^n_k] = R_\mathbf{D} \quad P_n \text{ a.s.},$$

as $|\xi_k| = 1,\ f(\cdots)^2 \in \mathbf{D}$ implies

$$|(\Delta M^f_k)^2| = |f(k,\mathbb{X})|^2 \in [r_\mathbf{D}, R_\mathbf{D}] \quad P_n \text{ a.s.}$$

□

**Proposition 2.3.11.** Let $\xi : \Omega \to \mathbb{R}$ be a continuous function satisfying $|\xi(\omega)| \leq a(1+ \parallel \omega \parallel_\infty)^b$ for some constants $a, b > 0$. Then,

$$\lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_{n/n}}} \mathbb{E}^\mathbb{Q}[\xi(\widehat{X}^n)] = \sup_{P\in\mathcal{Q}_\mathbf{D}} \mathbb{E}^P[\xi]. \tag{2.14}$$

*Proof.* To prove (2.14), we prove two separate inequalities together with a density argument which imply (2.14).

**First inequality** (for $\leq$ in (2.14))**:**

$$\limsup_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_{n/n}}} \mathbb{E}^\mathbb{Q}[\xi(\widehat{X}^n)] \leq \sup_{P\in\mathcal{Q}_\mathbf{D}} \mathbb{E}^P[\xi]. \tag{2.15}$$

By Lemma 2.3.10, we know that $\mathcal{Q}_{\mathbf{D}}^n \subseteq \mathcal{P}_{\mathbf{D}}^n$. For each $n \geq 1$, the inequality (2.7) implies that

$$\limsup_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \leq \sup_{P\in\mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi].$$

Since the convex hull of $\mathcal{Q}_{\mathbf{D}}$ is a weakly dense subset of $\mathcal{P}_{\mathbf{D}}$, see Proposition 2.3.8,

$$\limsup_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \leq \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi].$$

For all $n$, trivially $\sqrt{\mathbf{D}'_n/n} \subseteq \sqrt{\mathbf{D}/n}$ and $\mathcal{L}_n^{n+1} \subseteq \mathbb{R}^{n+1}$. Thus, $\mathcal{Q}^n_{\mathbf{D}'_n/n} \subseteq \mathcal{Q}^n_{\mathbf{D}/n}$. Hence, (2.15) follows.

**Second inequality** (for $\geq$ in (2.14))**:**

It remains to show that

$$\liminf_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \geq \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi].$$

For arbitrary $P \in \mathcal{Q}_{\mathbf{D}}$, we construct a sequence $(P^n)_n$ such that for all $n$,

$$P^n \in \mathcal{Q}^n_{\mathbf{D}'_n/n}, \tag{2.16}$$

and

$$\mathbb{E}^P[\xi] \leq \liminf_{n\to\infty} \mathbb{E}^{P^n}[\xi(\widehat{X}^n)]. \tag{2.17}$$

Fix $n$ and let $\xi_1, \ldots, \xi_n$ be some i.i.d sequence of random variables on $\Omega_n$ as defined in Section 2.2, i.e., $\xi_i : \Omega_n \to \{\pm 1\}$, for $i = 1, \ldots, n$. Now, we want to construct martingales $M^n$ whose laws are in $\mathcal{Q}^n_{\mathbf{D}'_n/n}$ and the laws of their interpolations tend to $P$. To achieve the above task, we introduce a scaled random walk with the piecewise constant càdlàg property (right continuity with left limits),

$$W_t^n := \frac{1}{\sqrt{n}} \sum_{l=1}^{\lfloor nt/T \rfloor} \xi_l = \frac{1}{\sqrt{n}} Z^n_{\lfloor nt/T \rfloor}, \quad 0 \leq t \leq T, \tag{2.18}$$

and we denote the continuous version of (2.18) obtained by linear interpolation by

$$\widehat{W}_t^n := \frac{1}{\sqrt{n}} \widehat{Z}^n_{\lfloor nt/T \rfloor}, \quad 0 \leq t \leq T. \tag{2.19}$$

By the central limit theorem;

$$(W^n, \widehat{W}^n) \Rightarrow (W, W)$$

as $n \to \infty$ on $D([0,T];\mathbb{R}^2)$ ($\Rightarrow$ implies convergence in distribution). i.e., the law $(P_n)$ converges to the law $P_0$ on the Skorohod space $D([0,T];\mathbb{R}^2)$ Billingsley [15, Theorem 27.1]. Let $g \in \mathcal{C}([0,T] \times \Omega, \sqrt{\mathbf{D}})$, such that

$$P = P_0 \circ \left( \underbrace{\int g(t,W)dW_t}_{M} \right)^{-1}.$$

Since $g$ is continuous and $\widehat{W}^n_t$ is the interpolated version of (2.18), it turns out that

$$\left( W^n, \left( g\left( \lfloor nt/T \rfloor T/n, \widehat{W}^n_t \right) \right)_{t\in[0,T]} \right) \Rightarrow \left( W, (g(t,W_t))_{t\in[0,T]} \right)$$

as $n \to \infty$ on $D([0,T];\mathbb{R}^2)$. We introduce martingales with discrete-time integrals,

$$M^n_k := \sum_{l=1}^{k} g\left( (l-1)T/n, \widehat{W}^n \right) \widehat{W}^n_{lT/n} - \widehat{W}^n_{(l-1)T/n}. \tag{2.20}$$

In order to construct a discretize martingale $M^n$ which is “close” to $M$ and also is such that $P_n \circ (M^n)^{-1} \in \mathcal{Q}^n_{\mathbf{D}'_n/n}$. We shall choose some

$$g_n : \{0, \ldots, n\} \times \mathcal{L}^{n+1}_n \to \sqrt{\mathbf{D}'_n/n},$$

such that,

$$M^n_k = \sum_{l=1}^{k} g_n\left( l-1, \frac{1}{\sqrt{n}} Z^n \right) \frac{1}{\sqrt{n}} \Delta Z^n_l.$$

Let $d_{J_1}$ be the Kolmogorov metric for the Skorohod $J_1$ topology. We choose $\widetilde{h}_n : \{0, \cdots, n\} \times \Omega \to \sqrt{\mathbf{D}'_n/n}$ such that

$$d_{J_1}\left( \left( \widetilde{h}_n(\lfloor nt/T \rfloor T/n, \widehat{W}^n_t) \right)_{t\in[0,T]}, \left( g(\lfloor nt/T \rfloor T/n, \widehat{W}^n_t) \right)_{t\in[0,T]} \right)$$

is minimal (this is possible because there are only finitely many choices for $\left( \widetilde{h}_n(\lfloor nt/T \rfloor T/n, \widehat{W}^n_t) \right)_{t\in[0,T]}$). This implies, due to the construction of $\mathbf{D}'_n$ as a discretization of $\mathbf{D}$ that

$$d_{J_1}\left(\left(\widetilde{h}_n(\lfloor nt/T\rfloor T/n,\widehat{W}^n_t)\right)_{t\in[0,T]},\left(g(\lfloor nt/T\rfloor T/n,\widehat{W}^n_t)\right)_{t\in[0,T]}\right)\to 0$$

as $n\to\infty$ on $D([0,T];\mathbb{R})$. From Billingsley [16, Theorem 3.1 and Theorem 14.1], it follows that

$$\left(W^n,\left(\widetilde{h}_n\left(\lfloor nt/T\rfloor T/n,\widehat{W}^n_t\right)\right)_{t\in[0,T]}\right)\Rightarrow\left(W,g(t,W_t)_{t\in[0,T]}\right)$$

as $n\to\infty$ on $D([0,T];\mathbb{R}^2)$. We then define $g_n:\{0,\ldots,n\}\times\mathcal{L}_n^{n+1}\to\sqrt{\mathbf{D}'_n/n}$ by

$$g_n:(\ell,\vec{\mathbb{X}})\mapsto\widetilde{h}_n(\ell,\widehat{\vec{\mathbb{X}}}).$$

Let $M^n$ be defined by

$$M^n_k=\sum_{l=1}^{k}g_n\left(l-1,\frac{1}{\sqrt{n}}Z^n\right)\frac{1}{\sqrt{n}}\Delta Z^n_l,\quad\forall k\in\{0,\cdots,n\}.$$

By stability of stochastic integral (see Duffie and Protter [40, Theorem 4.3 and Definition 4.1]),

$$\left(M^n_{\lfloor nt/T\rfloor}\right)_{t\in[0,T]}\Rightarrow M\quad\text{as } n\to\infty\text{ on } D([0,T];\mathbb{R})$$

because

$$M^n_{\lfloor nt/T\rfloor}=\sum_{l=1}^{\lfloor nt/T\rfloor}\widetilde{h}_n\left((l-1)T/n,\left(\widehat{W}_{kT/n}\right)_{k=0}^{n}\right)\Delta\widehat{W}_{lT/n}.$$

By Dolinsky et al. [38], the continuous version of (2.20) obtained by linear interpolation $\widehat{M}^n$ converges in distribution to $M$ on $\Omega$ endowed with the uniform metric on the Skorohod space, i.e., $\widehat{M}^n\Rightarrow M$ on $\Omega$. Since $\xi$ is bounded and continuous,

$$\lim_{n\to\infty}\mathbb{E}^{P_n\circ(M^n)^{-1}}[\xi(\widehat{X}^n)]=\mathbb{E}^{P_0\circ M^{-1}}[\xi].\tag{2.21}$$

Therefore, (2.16) is satisfied for $P^n=P_n\circ(M^n)^{-1}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}$. Trivially, (2.16) implies

$$\mathbb{E}^{P^n}[\xi(\widehat{X}^n)]\le\sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}}\mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)].\tag{2.22}$$

Combining (2.21) and (2.22), and taking the lim inf as $n$ tends to $\infty$, gives

$$\mathbb{E}^P[\xi] \leq \liminf_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)]. \tag{2.23}$$

Taking the supremum of (2.23) over $P \in \mathcal{Q}_{\mathbf{D}}$, the equation becomes

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] \leq \liminf_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)]. \tag{2.24}$$

Combining (2.15) and (2.24),

$$\begin{aligned}\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] &\geq \limsup_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \\ &\geq \liminf_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \\ &\geq \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi].\end{aligned}$$

Therefore,

$$\lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] = \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi]. \tag{2.25}$$

Density argument: (2.14) is established for all $\xi \in \mathcal{C}_b(\Omega, \mathbb{R})$ and also holds for all $\xi \in \mathbb{L}^1_*$ (see the density argument verification below). $\square$

**Proposition 2.3.12.** Let $\xi : \Omega \to \mathbb{R}$ be a continuous function satisfying $|\xi(\omega)| \leq a(1 + \|\omega\|_\infty)^b$ for some constants $a, b > 0$ and $\mathcal{Q}^n_{\mathbf{D}'_n}$ be the set of probability measures as defined in (2.4), then

$$\sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] = \max_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)]. \tag{2.26}$$

*Proof.* The left-hand side of (2.26) can be written as

$$\sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] = \sup_{f\in\mathcal{A}} \mathbb{E}^{P_n\circ(M^{f,\mathbb{X}})^{-1}}[\xi(\widehat{X}^n)],$$

where $\mathcal{A} = \left\{f : \{0, \ldots, n\} \times \mathcal{L}^{n+1}_n \to \sqrt{\mathbf{D}'_n}\right\}$ such that $f$ is $\mathcal{F}^n$-adapted. We shall prove that $\mathcal{A}$ is a compact subset of a finite-dimensional vector space, and that $f \mapsto \mathbb{E}^{P_n\circ(M^{f,\mathbb{X}})^{-1}}[\xi(\widehat{X}^n)]$ is continuous.

**First part**

Recall that for fixed $n \in \mathbb{N}$, $X^n = (X_k^n)_{k=0}^n$ is the canonical process defined by $X_k^n(x) = x_k$ for $x = (x_0, \dots, x_n) \in \mathcal{L}_n^{n+1}$, and $(\mathcal{F}_k^n)_{k=0}^n = \sigma(X_l^n, l = 0, \dots, k)$ is the filtration generated by $X^n$. We consider $\Omega_n = \{\omega = (\omega_1, \dots, \omega_n) : \omega_i \in \{\pm 1\},\ i = 1, \dots, n\}$ equipped with the power set. Let

$$P_n = \underbrace{\frac{\delta_{-1} + \delta_{+1}}{2} \otimes \cdots \otimes \frac{\delta_{-1} + \delta_{+1}}{2}}_{\text{n times}}$$

where for all $A \subseteq \mathbb{R}$,

$$\delta_x(A) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases},$$

be the product probability associated with the uniform distribution. $\xi_1, \dots, \xi_n$ is the i.i.d sequence of real-valued random variables such that $\xi_k$ belongs to $\{\pm 1\}$ and the components of $\xi_k$ are orthonormal in $L^2(P_n)$. We denote the associated random walk by $Z_k^n = \sum_{l=1}^k \xi_l$.

$\mathcal{A}$ is closed[1] and obviously bounded with respect to the norm $\|\cdot\|_\infty$ as $\mathbf{D}'_n$ is bounded[2]. By Heine-Borel theorem, $\mathcal{A}$ is a compact subset of a $N(n,n)$-dimensional vector space equipped with the norm $\|\cdot\|_\infty$.

**Second part**

Here, we want to show that $F : f \mapsto \mathbb{E}^{P_n \circ (M^{f,\mathbb{X}})^{-1}}[\xi(\widehat{X}^n)]$ is continuous.

$$\mathcal{Q}^n_{\mathbf{D}'_n} = \left\{ P_n \circ (M^{f,\mathbb{X}})^{-1};\ f : \{0, \dots, n\} \times \mathcal{L}_n^{n+1} \to \sqrt{\mathbf{D}'_n} \text{ is } \mathcal{F}^n\text{-adapted.} \right\}$$

where

$$M^{f,\mathbb{X}} = \left( \sum_{l=1}^k f(l-1, \mathbb{X}) \Delta \mathbb{X}_l \right)_{k=0}^n .$$

[1]The cardinality of $\mathcal{L}_n$, $\#\mathcal{L}_n = 2n+1$, $\#\mathcal{L}_n^{n+1} = (2n+1)^{n+1}$, and $\#(\{0, \dots, n\} \times \mathcal{L}_n^{n+1}) = (n+1)(2n+1)^{n+1} = N(n,n)$. Let $(f^m)_m \in \mathcal{A}^{N(n,n)}$ and $f : \{0, \dots, n\} \times \mathcal{L}_n^{n+1} \to \mathbb{R}$, such that $f^m \to f$, as $m \to \infty$, with respect to the maximum norm $\|\cdot\|_\infty$ (or any norm as a result of norm equivalency) on $\mathbb{R}^{N(n,n)}$. We have to prove that $f$ is adapted and $\sqrt{\mathbf{D}'_n}$-valued (is obvious, $\sqrt{\mathbf{D}'_n}$ is closed). For the first part, let $j \in \{0, \dots, n\}$. We want to show that $f(j, \cdot)$ is $\mathcal{F}_j^n$-measurable. This, however, follows from Billingsley [15, Theorem 13.4(*ii*)].

[2]If $V \in \mathbb{R}_{>0}$ such that $\mathbf{D}'_n \subseteq [0, V]$, then obviously $\|f\|_\infty = \max_{\substack{j \in \{0,\dots,n\} \\ \omega \in \mathcal{L}_n^{n+1}}} |f(j, \omega)| \le \sqrt{V}$.

$$
\begin{aligned}
\mathbb{E}^{P_n\circ(M^{f,\mathbb{X}})^{-1}}[\xi(\widehat{X}^n)] &= \int_{\mathcal{L}_n^{n+1}} \xi(\widehat{X}^n)dP_n\circ(M^{f,\mathbb{X}})^{-1}, \\
&= \int_{\Omega_n} \xi(\widehat{X}^n(M^{f,\mathbb{X}}))dP_n, \quad \text{(transforming measure)} \\
&= \sum_{\omega_n\in\Omega_n} P_n\{\omega_n\}\xi\circ(\widehat{X}^n)\circ M^{f,\mathbb{X}}(\omega_n).
\end{aligned}
$$

From Proposition 2.3.11 we know that $\xi$ is continuous, $\widehat{X}^n$ is the interpolated canonical process, i.e., $\widehat{X}:\mathcal{L}_n^{n+1}\to\Omega$, thus $\widehat{X}^n$ is continuous and $P_n$ takes it values from the set of real numbers. For $F: f\mapsto \mathbb{E}^{P_n\circ(M^{f,\mathbb{X}})^{-1}}[\xi(\widehat{X}^n)]$ to be continuous, $\psi: f\mapsto M^{f,\mathbb{X}}$ has to be continuous. Since $\mathcal{A}=\{f:\{0,\dots,n\}\times\mathcal{L}_n^{n+1}\to\sqrt{\mathbf{D}'}_n$, where $f$ is adapted with respect to the filtration generated by $\mathbb{X}\}$ is a compact subset of a $N(n,n)$-dimensional vector space for fixed $n\in\mathbb{N}$ and $M^{f,\mathbb{X}}:\Omega_n\to\mathcal{L}_n^{n+1}$, for all $f,g\in\mathcal{A}$,

$$|M^{f,\mathbb{X}}-M^{g,\mathbb{X}}| = |\|f\|_\infty-\|g\|_\infty| \le \|f-g\|_\infty.$$

Thus, $\psi$ is continuous with respect to the norm $\|\cdot\|_\infty$. Hence $F$ is continuous with respect to any norm[3] on $\mathbb{R}^{N(n,n)}$.

□

**Theorem 2.3.13.** Let $\xi:\Omega\to\mathbb{R}$ be a continuous function satisfying $|\xi(\omega)|\le a(1+\|\omega\|_\infty)^b$ for some constants $a,b>0$. Then,

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] = \lim_{n\to\infty} \max_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)]. \tag{2.27}$$

*Proof.* The proof follows directly from Proposition 2.3.11 and Proposition 2.3.12.

□

## Density argument verification

Let

$$f:\xi\mapsto \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi]$$

[3]For any two vector norms $\|\cdot\|_\alpha$, $\|\cdot\|_\beta$, and $C_1,C_2>0$, we have $C_1\|A\|_\alpha\le\|A\|_\beta\le C_2\|A\|_\alpha$, for all matrices $A\in\mathbb{R}^{N(n,n)}$. i.e., all norms on $\mathbb{R}^{N(n,n)}$ are equivalent because $\mathbb{R}^{N(n,n)}$ has $N(n,n)$-dimension for fixed $n\in\mathbb{N}$.

and

$$g : \xi \mapsto \lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_{n/n}}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)].$$

From (2.25), we know that for all $\xi \in \mathcal{C}_b(\Omega,\mathbb{R})$, $f(\xi) = g(\xi)$. Since $\mathbb{L}^1_*$ is the completion of $\mathcal{C}_b(\Omega,\mathbb{R})$ under the norm $\|\cdot\|_*$, $\mathcal{C}_b(\Omega,\mathbb{R})$ is dense in $\mathbb{L}^1_*$; and we want to prove for all $\xi \in \mathbb{L}^1_*$, $f(\xi) = g(\xi)$. To prove this, it is sufficient to show that $f$ and $g$ are continuous with respect to the norm $\|\cdot\|_*$.

**For continuity of $f$:**

For all $P \in \mathcal{Q}_{\mathbf{D}}$ and $\xi, \xi' \in \mathbb{L}^1_*$,

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] - \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi'] \le \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi - \xi']$$

and

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi - \xi'] \le \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[|\xi - \xi'|].$$

Since, $\mathcal{Q}_{\mathbf{D}} \subseteq \mathcal{Q}$,

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[|\xi - \xi'|] \le \sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[|\xi - \xi'|] = \|\xi - \xi'\|_*.$$

Then,

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] - \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi'] \le \|\xi - \xi'\|_*. \tag{2.28}$$

Interchanging $\xi$ and $\xi'$,

$$\sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi'] - \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] \le \|\xi' - \xi\|_*. \tag{2.29}$$

Adding (2.28) and (2.29), we have

$$\left| \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi] - \sup_{P\in\mathcal{Q}_{\mathbf{D}}} \mathbb{E}^P[\xi'] \right| \le \|\xi - \xi'\|_*. \tag{2.30}$$

Hence,

$$|f(\xi) - f(\xi')| \le \|\xi - \xi'\|_*.$$

**For continuity of $g$:**

We can follow the same argument as above; for all $\mathbb{Q} \in \mathcal{Q}^n_{\mathbf{D}'_n/n}$, $\xi, \xi^{'} \in \mathbb{L}^1_*$ and for all $n$,

$$\begin{aligned}\sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] - \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi^{'}(\widehat{X}^n)] \\ \leq \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n) - \xi^{'}(\widehat{X}^n)] \\ \leq \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[|\xi(\widehat{X}^n) - \xi^{'}(\widehat{X}^n)|].\end{aligned}$$

Since, $\mathcal{Q}^n_{\mathbf{D}'_n/n} \subseteq \mathcal{Q}^n_{\mathbf{D}/n}$ and $\mathcal{Q}^n_{\mathbf{D}/n} \subseteq \mathcal{Q}$, we can say that

$$\sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[|\xi(\widehat{X}^n) - \xi^{'}(\widehat{X}^n)|] \leq \sup_{Q\in\mathcal{Q}} \mathbb{E}^{Q}[|\xi - \xi^{'}|] = \|\xi - \xi^{'}\|_*,$$

then,

$$\sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] - \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi^{'}(\widehat{X}^n)] \leq \|\xi - \xi^{'}\|_*. \tag{2.31}$$

Taking the limit when $n$ goes to $\infty$, (2.31) becomes,

$$\lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] - \lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi^{'}(\widehat{X}^n)] \leq \|\xi - \xi^{'}\|_*. \tag{2.32}$$

Interchanging $\xi$ and $\xi^{'}$,

$$\lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi^{'}(\widehat{X}^n)] - \lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \leq \|\xi^{'} - \xi\|_*. \tag{2.33}$$

Adding (2.32) and (2.33), we have

$$\left| \lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] - \lim_{n\to\infty} \sup_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n/n}} \mathbb{E}^{\mathbb{Q}}[\xi^{'}(\widehat{X}^n)] \right| \leq \|\xi - \xi^{'}\|_*.$$

Hence,

$$\left| g(\xi) - g(\xi^{'}) \right| \leq \|\xi - \xi^{'}\|_*.$$

# Chapter 3

# A Simplified Approach to Nonstandard Measure Theory

## 3.1 Introduction

We give a simplified introduction to nonstandard measure theory that does not presuppose prior acquaintance with mathematical logic. This approach requires no previous knowledge of nonstandard analysis. The methodology is presented in terms of sequences, equivalence relations and equivalence classes with respect to binary measures. The approach is based on Lindstrøm's [73] work. However, our approach is more simplified. We construct the extended nonstandard enlargement in measure theoretic language and also show how the language of logic relates to the mathematical discourse in probability theory. We provide an alternative construction of the renowned Loeb measure using basic knowledge of real analysis.

Nonstandard analysis, or the theory of infinitesimals as some people prefer to call it (see Robinson [94]), was introduced by Abraham Robinson in 1961. Robinson in [94] introduced a mathematical foundation for infinitesimals. His work started with a mathematical object such as the system of real numbers or some Banach space. Being a mathematical logician, he used a formal language to express facts about the mathematical structure he was working on, and called the structure a *standard model* for the true statement expressed in the formal language. One can think of the standard model as a universe in some sense, thus, it is also known as the *standard universe* Hrbáček [55]. Robinson went further to prove that there also exists another mathematical structure called the *nonstandard model* such that in the standard and the nonstandard models the same first-order statements are true, and the nonstandard model for the system of real numbers may be constructed in the form of ultrapowers as defined by Łoś [78]. This nonstandard model contains infinitely small and infinitely large numbers in a well-defined sense.

One way to summarize Robinson's [94] result is as follows: We have two universes, the standard universe and the nonstandard universe, and every first-order statement that is true in the standard universe is also true in the nonstandard universe and the standard model can be embedded into the nonstandard model. This implies that one can use the nonstandard model to analyze the standard model, thus, Robinson called his method and results *nonstandard analysis*.

Many authors have introduced a more elementary and intuitive approach to nonstandard analysis. Schmieden and Laugwitz [98] introduced a nonstandard analysis that is more constructive, however their approach is much heavily based on classical analysis without the use of model theory (mathematical logic). The work of Schmieden and Laugwitz [98] was further developed in subsequent papers by Laugwitz [66, 67, 68]. In their approach, one of the important properties of nonstandard analysis, the Transfer Principle (see Theorem 3.4.5), is lacking. Thus, their nonstandard model cannot be related to the standard model as easily as in Robinson's result. Mycielski [83] introduced a locally constructive theory of infinitesimals. In his theory, every proof can be interpreted in a finite model. Lindstrøm's [73] presentation of Robinson's approach to nonstandard analysis is far more intuitive and user friendly. His approach is presented in terms of sequences, equivalence relation and equivalence classes with respect to binary measures. There are other interesting approaches in the literature, see, for example, Keisler [59]. Our approach to nonstandard analysis is based on Lindstrøm's [73] work. We present the nonstandard analysis in a language that is easily accessible to a wider audience of mathematicians rather than only logicians.

One of the important features of nonstandard analysis is that it allows one to embed a standard mathematical theory for instance, measure theory, probability theory or algebra, in a standard universe into the nonstandard universe. Most results in the nonstandard universe are profound and easy to follow, for instance, there are several existence theorems whose only known proofs use nonstandard analysis. For example, Perkins [90] proved a global characterization of (standard) Brownian local time with nonstandard analysis. Anderson and Raimondo [5] use nonstandard analysis to prove the existence of equilibrium in a continuous time financial model where the number of underlying assets is at least one more than the number of independent source of risk.

The embeddment of a standard universe into a nonstandard universe enables one to establish a two way interaction between a standard universe and a nonstandard universe in the sense that a statement is true in the standard universe if and only if the *transfer* of the statement into the nonstandard universe is true. This mechanism is basically called the *Transfer Principle*. However, the strength of nonstandard analysis is not limited to the Transfer Principle. As a matter of fact, transfer is limited between the *standard entities* in the standard universe and the so-called *internal entities* in the nonstandard universe. The nonstandard universe that contains only these internal entities is called the *internal universe*.

Another important property of nonstandard analysis is the *Internal Definition Principle* (see Theorem 3.4.8), which tells that the subsets of internal sets in the internal universe are always internal. The Internal Definition Principle is a consequence of the Transfer Principle.

The *Saturation Principle* (see Theorem 3.4.10), more precisely the Countable Saturation Principle, is almost as important as the Transfer Principle and the Internal Definition Principle. It asserts that the intersection of any decreasing countable sequence of nonempty elements of the internal universe is always nonempty. A nonstandard universe that is adequate for applications especially when the target space satisfies some countability condition will satisfy the Countable Saturated Principle. As an application of the Countable Saturated Principle, we discuss the *Loeb space*, Loeb [76].

The Loeb space can be "seen" as the standard reduct of a nonstandard probability space. The Countable Saturation Principle ensures that the measure (Loeb measure) defined on the Loeb space is countably additive. The richness of the Loeb measures makes them applicable in a wide range of research: stochastic analysis Müller [82], Anderson [3], Cutland et al. [31], financial economics Cutland et al. [30], Duffie and Protter [40], control theory Berg [12], mathematical physics Albeverio et al. [2], (for recent overview, see Cutland [29]).

This chapter is organized as follows: In Section 3.2, we explain the construction of the *hyperreals* by extending the real line to accommodate the infinitely small and the infinitely large numbers. We begin Section 3.3 by defining the standard universe (*superstructure*) and its components and then show how the language of logic relates to the mathematical discourse in probability theory. Section 3.4 explains the construction of the nonstandard universe (the image of the superstructure under the nonstandard embedding $^*$). Thereafter, we explain some basic properties of the nonstandard enlargement. In Section 3.5, we provide an alternative construction of the Loeb measure using basic knowledge of real analysis. We shall do this by establishing that an internal finitely additive measure induces a premeasure and then extend the premeasure into a measure on the $\sigma$-algebra. We conclude by proving the uniqueness of this measure.

## 3.2 Preliminaries

We begin by extending the real line to contain both the infinitely small and infinitely large. Recall: One way to construct the real line is to add new rational points to represent limits of convergence of rationals. In this approach one has to identify sequences converging to the same point in $\mathbb{R}$. This runs as follows: Let $\equiv$ be the *equivalence relation* on the set $\mathbb{S}$ of all rational Cauchy sequences defined by

$$(f_n)_n \equiv (g_n)_n \iff \lim_{n\to\infty}(f_n - g_n) = 0.$$

Then the reals are the set

$$\mathbb{R} = \mathbb{S}/\equiv$$

of all equivalence classes. In order to construct $^*\mathbb{R}$ (hyperreals) from $\mathbb{R}$, a well-organized structure that does not entail only the limit of convergence but also the mode of convergence is required. To achieve this, one needs to identify as few sequences as possible. i.e., the *trivial identification*, (cf. Lindstrøm [73]);

$$(f_n)_n \sim (g_n)_n \iff (f_n) = (g_n),$$

where $\sim$ is the equivalence relation. But, if $\mathbf{f} = (f_n)_n$ is a sequence such that $f_n = 0$ if and only if $n$ is even and $\mathbf{g} = (g_n)_n$ is a sequence such that $g_n = 0$ if and only if $n$ is odd; then $\mathbf{f} \cdot \mathbf{g} = 0$, although, both $f$ and $g$ are non-zero. Thus, the trivial identification gives rise to a structure with zero divisors (cf. Lindstrøm [73]). The task now is to make the equivalence relation $\sim$ strong enough to avoid the problem of zero divisors. Thus, we have to fix a finitely additive measure on $\mathbb{N}$ with the following properties.

**Definition 3.2.1.** (cf. Lindstrøm [73, Definition 1.1.2 ]) Let $\mu$ be a $\{0,1\}$-valued finitely additive measure on the set $\mathbb{N}$ of positive integers such that:

$(a)$ $\mu(E)$ is defined and it is either 1 or 0 for all $E \subset \mathbb{N}$.

$(b)$ If $\mu(E) = 1$ and $\mu(F) = 1$, then $\mu(E \cap F) = 1$ for all $E, F \subset \mathbb{N}$.

$(c)$ $\mu(\mathbb{N}) = 1$ and $\mu(E) = 0$ for all finite $E$.

$(d)$ For any $E \subset \mathbb{N}$, either $\mu(E) = 1$ or $\mu(E^c) = 1$, (but not both).

The measure $\mu$ is a finitely additive measure means $\mu(E \cup F) = \mu(E) + \mu(F)$ for all disjoint sets $E$ and $F$. It is important to note that $\mu$ divides the subsets of $\mathbb{N}$ into two different classes; the "large ones" with measure one and the "small ones" with measure zero, such that all finite sets are small because they have measure zero (cf. Lindstrøm [73]).

**Remark 3.2.2.** If $\mu(E) = 1$ and $E \subseteq F \subset \mathbb{N}$, then $\mu(F) = 1$.

**Definition 3.2.3.** (cf. Lindstrøm [73, Definition 1.1.3]) Let $\mathbb{R}^{\mathbb{N}}$ (the direct product of $\mathbb{N}$ copies of $\mathbb{R}$) be the set of all sequences of real numbers and let $\sim$ be the equivalence relation on $\mathbb{R}^{\mathbb{N}}$ defined by

$$(f_n)_n \sim (g_n)_n \iff \mu\{n : f_n = g_n\} = 1.$$

That is $(f_n)_n$ equal to $(g_n)_n$ almost everywhere.

**Definition 3.2.4.** Let $\langle (f_n)_n \rangle$ denote the *equivalence class* of the sequence $(f_n)_n$ in $\mathbb{R}^{\mathbb{N}}$. Addition, multiplication and absolute value (norm) is defined componentwise by

$$\langle (f_n)_n \rangle + \langle (g_n)_n \rangle := \langle (f_n + g_n)_n \rangle; \quad \langle (f_n)_n \rangle \cdot \langle (g_n)_n \rangle := \langle (f_n \cdot g_n)_n \rangle,$$

and $|\langle (f_n)_n \rangle| := \langle (|f_n|)_n \rangle$ respectively.

**Definition 3.2.5.** Let $\mathbb{R}^{\mathbb{N}}$ be the set of all sequences of real numbers and let $\sim$ be the equivalence relation. The *hyperreal* is given by

$$^*\mathbb{R} = \mathbb{R}^{\mathbb{N}} / \sim .$$

Thus, for every sequence $(f_n)_n$ in $\mathbb{R}^{\mathbb{N}}$, $\langle (f_n)_n \rangle$ denotes its image in $^*\mathbb{R}$. We then have a natural embedding

$$^* : \mathbb{R} \to {}^*\mathbb{R}$$

by taking ${}^*f = \langle f \rangle$.

It is possible to extend the operations and relations of $\mathbb{R}$ to ${}^*\mathbb{R}$: for arbitrary $\langle f \rangle = \langle (f_n)_n \rangle$ and $\langle g \rangle = \langle (g_n)_n \rangle$ in ${}^*\mathbb{R}$,

$$\langle f \rangle + \langle g \rangle = \langle h \rangle \iff \mu\{n : f_n + g_n = h_n\} = 1$$

and

$$\langle f \rangle \cdot \langle g \rangle = \langle h \rangle \iff \mu\{n : f_n \cdot g_n = h_n\} = 1.$$

In a similar manner,

$$\langle f \rangle < \langle g \rangle \iff \mu\{n : f_n < g_n\} = 1. \tag{3.1}$$

By the definition of $<$ in ${}^*\mathbb{R}$, one can easily see that ${}^*\mathbb{R}$ is linearly ordered. For example, let us prove the transitivity of ${}^*\mathbb{R}$.

*Proof.* Let $\langle f \rangle < \langle g \rangle$ and $\langle g \rangle < \langle h \rangle$. We want to show that $\langle f \rangle < \langle h \rangle$. From (3.1), we know that $\mu\{n : f_n < g_n\} = 1$ and $\mu\{n : g_n < h_n\} = 1$. It follows from Definition 3.2.1-(b) that

$$\mu\{n : f_n < g_n \cap g_n < h_n\} = 1.$$

Thus, $f_n < g_n$ and $g_n < h_n$, and by transitivity of $<$ in $\mathbb{R}$, $f_n < h_n$.
Hence,

$$\mu\{n : f_n < h_n\} = 1.$$

□

By Definition 3.2.5, in ${}^*\mathbb{R}$, either the sequence $\mathbf{f}$, where $f_n = 0$ if and only if $n$ is even, or the sequence $\mathbf{g}$ where $g_n = 0$ if and only if $n$ is odd, will be identified with the zero sequence $\mathbf{0}$ and the other one with $\mathbf{1}$ (cf. Lindstrøm [73]).

We identify $\gamma \in \mathbb{R}$ with $\langle (\gamma)_n \rangle \in {}^*\mathbb{R}$.

**Definition 3.2.6.** For every sequence $(f_n)_n$ in $\mathbb{R}^{\mathbb{N}}$ and $\langle f \rangle = \langle (f_n)_n \rangle$ in ${}^*\mathbb{R}$;

($a$) We say that $\langle f \rangle$ is *infinitesimal* and write $\langle f \rangle \simeq 0$ if and only if

$$\forall r \in \mathbb{R}_{>0},\ \mu\{n : |f_n| \leq r\} = 1. \tag{3.2}$$

($b$) $\langle f\rangle$ is *limited* if and only if

$$\exists r \in \mathbb{R}_{>0},\ \mu\{n : -r < f_n < r\} = 1. \tag{3.3}$$

($c$) For arbitrary $\langle g\rangle$ in ${}^*\mathbb{R}$, $\langle f\rangle$ is *infinitely close* to $\langle g\rangle$ if and only if $\langle |f-g|\rangle$ is infinitesimal:

$$\langle f\rangle \simeq \langle g\rangle \iff \langle |f-g|\rangle \simeq 0. \tag{3.4}$$

($d$) $\langle f\rangle$ is *unlimited* if and only if $\langle f\rangle$ is not limited.

## 3.3 The standard enlargement and its components

The domain of reals is not large enough for development of contemporary mathematics most especially in measure theory and probability theory. Thus, the nonstandard extension of reals is not sufficient for a fully fledged application of nonstandard analysis to problems in measure theory and probability theory. In order to have an effective nonstandard framework, one needs an extended universe that contains not only numbers and functions, but also mathematical objects such as sets of functions, sets of spaces of functions, topological spaces, measure spaces etc. In view of this, we introduce the superstructure (see below) over $K$. One can assume $K$ is large enough to contain all mathematical objects. These objects can be defined as sets in the superstructure of $K$. Every object of standard mathematics lives in the superstructure.

**Definition 3.3.1.** For any set $K$ (where $K$ is regarded as a set of *individuals*, i.e., if $x \in K$ then $x$ has no elements.), $V(K)$ is a *superstructure* if

$$V_0(K) = K, \quad V_{k+1}(K) = V_k(K) \cup \mathcal{P}(V_k(K)), \quad \text{and} \quad V(K) = \bigcup_{k\in\mathbb{N}} V_k(K),$$

where $\mathcal{P}(A)$ denote the power set of $A$.

The elements of this superstructure are precisely the mathematical objects that can be obtained by iterating the power set operator countably many times. For every object $a$ in $V(K)$, $a$ is either an element in $K$ or a set that belongs to $V(K)\backslash K$. The *rank* of an object $a$ in $V(K)$ is the smallest $k$ for which $a$ is in $V_k(K)$. It is important to note that $K = V_0(K) \subset V_1(K) \subset \cdots$, and $K = V_0(K) \in V_1(K) \in \cdots$. Thus, when $i$ is less than $j$, $V_i(K)$ becomes an element in $V_j(K)$, and objects with rank greater than or equal to 1 in $V(K)$ are precisely the sets in $V(K)$. The objects in $V_0(K)$ have rank 0 and the empty set $\emptyset$ has rank 1. If $a$ is an object in $V(K)$ with rank greater than 1, and $b$ is an element in $a$, then $b$ is also an object in $V(K)$ and the rank of $b$ is strictly less than the rank of $a$.

**Definition 3.3.2.** Let $\mathcal{L}_{V(K)}$ be the *language* of $V(K)$, having the set of symbols given by $\{\dot{v}_n : n \in \mathbb{N}\} \cup \{\dot{a} : a \in V(K)\} \cup \{\dot{\in}, \dot{=}, \dot{\wedge}, \dot{\neg}, \dot{\forall}, \dot{\exists}, \dot{(}, \dot{)}\}$, in which

($a$) $\dot{v}_1, \dot{v}_2, \dot{v}_3, \dot{v}_4 \cdots$ are variables;

($b$) $\dot{a}$ is a constant symbol for each $a \in V(K)$;

($c$) $\dot{\in}, \dot{=}$ are relation symbols;

($d$) $\dot{\wedge}$ (and), $\dot{\neg}$ (not) are connectives;

($e$) $\dot{\forall}, \dot{\exists}$ are bounded quantifiers;

($f$) $\dot{(}, \dot{)}$ are parentheses.

A *string* of $\mathcal{L}_{V(K)}$ is a finite sequence of symbols of $\mathcal{L}_{V(K)}$. A string is an *atomic formula* if and only if it is of the form $\dot{v}_k \dot{=} \dot{v}_l$ or $\dot{v}_k \dot{\in} \dot{v}_l$ where $\dot{v}_k$ and $\dot{v}_l$ are variables for some $k, l \in \mathbb{N}$. The set $\mathbf{F}$ of all formulas is the smallest subset of strings which contains all the atomic formulas, and for all $\psi, \phi \in \mathbf{F}$ and any variable $\dot{v}_k$, the strings $\dot{(}\dot{\exists}\dot{v}_k \dot{\in} \dot{v}_l\dot{)}\psi$, $\dot{(}\dot{\forall}\dot{v}_k \dot{\in} \dot{v}_l\dot{)}\psi$, $\dot{(}\dot{\exists}\dot{v}_k \dot{\in} \dot{a}\dot{)}\psi$, $\dot{(}\dot{\forall}\dot{v}_k \dot{\in} \dot{a}\dot{)}\psi$, $\dot{\neg}\psi$, $\dot{(}\psi\dot{\wedge}\phi\dot{)}$ are all in $\mathbf{F}$.

For any variable $\dot{v}_k$, $\dot{v}_k$-*quantifier* means the string $\dot{(}\dot{\exists}\dot{v}_k \dot{\in} \dot{a}\dot{)}$. When we refer to the specific occurrence of the $\dot{v}_k$-quantifier of a given formula, we *underline* its position in the formula. For every formula $\psi$, strings $X, Y$, and variable $\dot{v}_l$, the *scope* of an occurrence of the $\dot{v}_k$-quantifier in a formula is defined recursively as follows:

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

- The scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}$ in $\psi\dot{\wedge}X\underline{\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}}Y$ also equals the scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}$ in $X\underline{\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}}Y$.

- The scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}$ in $\dot{(}\dot{\exists}\dot{v}_l\dot{)}X\underline{\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}}Y$ equals the scope of the occurrence of $\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}$ in $X\underline{\dot{(}\dot{\exists}\dot{v}_k\dot{\in}\dot{a}\dot{)}}Y$.

The occurrence of a variable $\dot{v}_k$ in a formula is called a *bound* if and only if it occurs in the scope of a $\dot{v}_k$-quantifier in the formula. Otherwise, the occurrence of $\dot{v}_k$ is said to be *free*. Thus, any particular occurrence is either free or bound, but not both. However, a variable $\dot{v}_k$ can have both free and bound occurrence in the same formula. For a formula $\psi$, if the occurrence of $\dot{v}_k$ in $\psi$ is free, then we often denote it as $\psi(\dot{v}_k)$. And for any given $\psi$ and any occurrence of $\dot{v}_k$ in $\psi$, one can tell if the occurrence of $\dot{v}_k$ in $\psi$ is free or bound by how $\psi$ is constructed from the atomic formulas. A formula $\psi$ is called a *sentence* if and only if all the occurrences of the variables in $\psi$ are not free, Bell and Slomson [11].
Let $\{\dot{v}_n\}_{n\in\mathbb{N}}$ be the sequence of all variables. An interpretation is a map

$$I : \{\dot{v}_n\}_{n\in\mathbb{N}} \cup \{\dot{a} : a \in V(K)\} \longrightarrow V(K)$$

such that $I(a) = a$ for all $a \in V(K)$ and a map

$$\alpha(\ \cdot\ |V(K), I) : \mathbf{F} \to \{0, 1\}.$$

The "interpretation" of a formula is the assignments of the *truth values* true or false to the formula, relative to the interpretation (valuation) $I$ of the variables. Ultimately having recursively defined $\alpha$ (see below), we shall define the $\models$ relation as follows:

$$V(K) \models_I \psi \iff \alpha(\psi|V(K), I) = 1.$$

The truth of a given formula is defined in terms of the components of the formula as follows:

1. Suppose $\dot{v}_k\dot{\in}\dot{a}$ is a formula.

$$\alpha(\dot{v}_k\dot{\in}\dot{a}|V(K), I) = \begin{cases} 1 & \text{if, } I(\dot{v}_k) \in I(\dot{a}) \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

2. Suppose $\dot{v}_k \dot{=} \dot{v}_l$ is a formula.

$$\alpha(\dot{v}_k \dot{=} \dot{v}_l | V(K), I) = \begin{cases} 1 & \text{if, } I(\dot{v}_k) = I(\dot{v}_l) \\ 0 & \text{otherwise.} \end{cases} \tag{3.6}$$

3. Suppose $\psi \dot{\wedge} \phi$ is a formula. $\alpha(\psi \dot{\wedge} \phi | V(K), I) = 1$ if and only if $\alpha(\psi | V(K), I) = 1$ and $\alpha(\phi | V(K), I) = 1$. Thus,

$$\alpha(\psi \dot{\wedge} \phi | V(K), I) = \min\{\alpha(\psi | V(K), I), \alpha(\phi | V(K), I)\} \tag{3.7}$$

4. Suppose $\dot{\neg}\psi$ is a formula. $\alpha(\dot{\neg}\psi | V(K), I) = 1$ if and only if $\alpha(\psi | V(K), I) = 0$. Thus,

$$\alpha(\neg\psi | V(K), I) = 1 - \alpha(\psi | V(K), I). \tag{3.8}$$

5. Let $(\dot{\exists} \dot{v}_k \dot{\in} \dot{a})\psi$ be a formula where $\dot{\exists} \dot{v}_k \dot{\in} \dot{a}$ does not occur in $\psi$ for any variable $\dot{v}_k$ and constant $\dot{a}$. We denotes $J$ as an interpretation.

$$\alpha\Big((\dot{\exists} \dot{v}_k \dot{\in} \dot{a})\psi | V(K), I\Big) = \max_{\dot{v}_k} \left\{ \begin{array}{l} \alpha\big(\psi | V(K), J\big) : J(\dot{v}_k) \in I(\dot{a}) \\ \text{ and for all } x \in \{\dot{v}_n : n \neq k\} \cup \{\dot{a} : a \in V(K)\}, \\ J(x) = I(x). \end{array} \right\}. \tag{3.9}$$

### 3.3.1 Link between mathematical logic and probability theory

It is important to note that formal expressions of these forms (as given above) are frequently use in probability theory. Recall some elements of probability theory. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where $\Omega$ is the set of all possible event, $\mathcal{F}$ is a $\sigma$-algebra, and $\mathbb{P}$ is the probability measure that assigns a probability $\mathbb{P}(B)$ to every event $B \in \mathcal{F}$ such that $\mathbb{P} : \mathcal{F} \to [0, 1]$. Let $X$ and $Y$ be real-valued random variables and let $(X_n)_n$ be the sequence of real-valued random variables. The following expressions in probability theory are treated as objects of mathematical discourse:

1. Relation symbols $(\dot{\in}, \dot{=})$:

- Consider $\mathbb{P}\{X \in A\}$. $X \in A$ is a well defined expression in probability theory. The assignments of the truth values to the formula $X\dot{\in}A$ is given in (3.5).

- Consider $\mathbb{P}\{X = Y\}$. $X = Y$ is a well defined expression in probability theory. Thus, the assignments of the truth values to the formula $X\dot{=}Y$ is given in (3.6).

2. Connectives ($\dot{\wedge}$, $\dot{\neg}$):

   - Consider $\mathbb{P}\{X \in [0,1]\&Y \geq 0\}$. $X \in [0,1]\&Y \geq 0$ is a well defined expression in probability theory. The assignments of the truth values to the formula "$\psi = X \in [0,1]$ and $\phi = Y \geq 0$" is given in (3.7).
   - Recall $\mathbb{P}\{X \in A\}$. The assignments of the truth values to the formula " not $\psi$", where $\psi = X\dot{\in}A$ is given in (3.8).

3. Quantifier ($\exists$): Consider $\mathbb{P}\{\exists n \in \mathbb{N}, X_n \geq 0\}$. $\exists n \in \mathbb{N}, X_n \geq 0$ is a well defined expression in probability theory. The assignments of the truth values to the formula $(\exists\dot{v}_k\dot{\in}\dot{a})\psi$, assuming $\exists\dot{v}_k\dot{\in}\dot{a}$ does not occur in $\psi$, is given in (3.9).

Other existing logical symbols, for example, disjunction $\vee$, implication $\rightarrow$, equivalence $\leftrightarrow$ and universal quantifier $\forall$, can be abbreviated using the above cases. They are used in the following ways:

1. $\psi\dot{\vee}\phi$ abbreviates $\dot{\neg}\dot{(}\dot{\neg}\psi\dot{\wedge}\dot{\neg}\phi\dot{)}$.
2. $\psi\dot{\rightarrow}\phi$ abbreviates $\dot{\neg}\dot{(}\psi\dot{\wedge}\dot{\neg}\phi\dot{)}$.
3. $\psi\dot{\leftrightarrow}\phi$ abbreviates $\dot{(}\psi\dot{\rightarrow}\phi\dot{)}\dot{\wedge}\dot{(}\phi\dot{\rightarrow}\psi\dot{)}$.
4. $\dot{(}\dot{\forall}\dot{v}_k \in \dot{a}\dot{)}\psi = \dot{\neg}\dot{(}\dot{\exists}\dot{v}_k \in \dot{a}\dot{)}\dot{\neg}\psi$.

**Lemma 3.3.3.** If $\psi$ contains no free variables, then for all interpretations $I$ and $J$,

$$V(K) \models_I \psi \iff V(K) \models_J \psi.$$

*Proof.* The proof is an induction on the complexity of $\psi$. For all $k, l \in \mathbb{N}$, let $\dot{v}_k, \dot{v}_l$ be variables of an arbitrary set in $V(K)$. Then

$$V(K) \models_I \dot{v}_k = \dot{v}_l \iff I(\dot{v}_k) = I(\dot{v}_l).$$

Suppose $\psi$ is a formula. For every $k \in \mathbb{N}$, if $\dot{v}_k$ occur free in $\psi$, then the truth or falsity of

$$V(K) \models_I \psi$$

depends only on the value of $I(\dot{v}_k)$ for every variable $\dot{v}_k$ that is free in $\psi$. That is, for all interpretations $I$ and $J$, $I(\dot{v}_k) = J(\dot{v}_k)$. Then,

$$V(K) \models_I \psi \iff V(K) \models_J \psi.$$

If $\psi$ is a sentence, then the truth or falsity of

$$V(K) \models_I \psi$$

is completely independent of $I$ because $\psi$ (as a sentence) contains no free variables. Thus, for all interpretations $I$ and $J$, we may suppress the interpretation and write

$$V(K) \models \psi.$$

$\square$

**Notation 3.3.4.** We drop brackets and even dots where no notation ambiguity can arise.

## 3.4 Construction and some properties of the nonstandard enlargement

Here we explain the setting of nonstandard analysis introduced in Robinson and Zakon [95]. The construction of the image of the classical superstructure under the nonstandard embedding $^*$ proceeds in two stages. Firstly, we construct a *bounded ultrapower* of $V(K)$ using bounded sequences of elements in $V(K)$ with measure one. Secondly, we map the bounded ultrapower into the superstructure $V(^*K)$ in such a way that the embedding satisfies the Transfer Principle.

A sequence $(A_n)_n$ is *rank bounded* if there is a fixed $k \in \mathbb{N}$ such that $A_n \in V_k(K)$ for all $n$. If $(A_n)_n$ is bounded,

$$\mathbb{N} = \{n : A_n \text{ has rank } 0\} \cup \cdots \cup \{n : A_n \text{ has rank } k\}$$

and

$$\mu\{n : A_n \text{ has rank } i\} = \begin{cases} 1 & \text{if, } i \leq k \\ 0 & \text{if, } i > k, \end{cases}$$

thus, the rank of $(A_n)_n$ is $k$. Let $A = (A_n)_n$ and $B = (B_n)_n$ be two bounded sequences. $A$ and $B$ are *equivalent* if and only if $(A_n)_n = (B_n)_n$ $\mu$ almost surely:

$$A \sim^\mu B \iff \Big(\mu\{n : V(K) \models A_n \dot{=} B_n\} = 1\Big).$$

Let $A^\mu$ denote the *equivalence class* of a bounded sequence $A$ and define the set of all equivalence classes as

$$V(K)^{\mathbb{N}}/\sim^\mu = \Big\{A^\mu : (\exists m \in \mathbb{N})\Big(\mu\{n : V(K) \models A_n \dot{\in} V_m(K)\} = 1\Big)\Big\}.$$

(We assume $K \cap \mathbb{N} = \emptyset$; though, there may be a copy of $\mathbb{N}$ in K, for example, if $\mathbb{R}$ is defined as a set of the equivalence classes of the sequence of rational numbers, $K = \mathbb{R}$). By definition, all the elements of $V(K)^{\mathbb{N}}/\sim^\mu$ are bounded with respect to the rank of the superstructure with measure one. $V(K)^{\mathbb{N}}/\sim^\mu$ is called the *bounded ultrapower* of $V(K)$.

**Claim 1.** $B^{'} \in A^{'} \in V_{k+1}(K)$ implies $B^{'} \in V_k(K)$.

*Proof of Claim 1.* By definition of a superstructure, $A^{'} \in V_{k+1}(K)$ simply implies $A^{'} \in V_k(K)$ or $A^{'} \subseteq V_k(K)$, and objects of $V_k(K)$ are either elements of $K$ or sets of $V_k(K) \setminus K$. Thus, $A^{'} \in V_{k+1}(K)$ is either $A^{'} \in K$ or $A^{'} \subseteq V_k(K)$. If $B^{'} \in A^{'}$, then $A^{'} \notin K$ ($K$ is a set of individuals). Hence, $B^{'} \in A' \in V_{k+1}(K)$ implies $B^{'} \in A' \subseteq V_k(K)$, that is, $B^{'} \in V_k(K)$. Since $V_k(K) \in V_{k+1}(K) \subset V(K)$, each $V_k(K)$ is an element of $V(K)$. □

The *membership relation* $\in^\mu$ on $V(K)^{\mathbb{N}}/\sim^\mu$ is defined as follows:

$$\in^\mu := \Big\{(B^\mu, A^\mu) : \Big(\mu\{n : V(K) \models B_n \dot{\in} A_n\} = 1\Big)\Big\}.$$

That is,

$$B^\mu \in^\mu A^\mu \iff \Big(\mu\{n : V(K) \models B_n \dot{\in} A_n\} = 1\Big).$$

Thus, there exist a *canonical embedding* of $A \mapsto A^\mu$ and a *natural proper embedding*

$$i : V(K) \longrightarrow V(K)^{\mathbb{N}}/\sim^\mu$$

where $i(A)$ is the equivalence class corresponding to the constant sequence $A$. Let ${}^*K = K^{\mathbb{N}}/\sim^{\mu}$. We need to construct an injective map

$$j : V(K)^{\mathbb{N}}/\sim^{\mu} \longrightarrow V({}^*K).$$

By definition, the bounded ultrapower is the union of the chain $V_0(K)^{\mathbb{N}}/\sim^{\mu} \subseteq \cdots \subseteq V_k(K)^{\mathbb{N}}/\sim^{\mu} \subseteq \cdots$ and we can define $j$ by recursion. For $k = 0$, $j$ must be the identity on ${}^*K$. i.e.,

$$j : x \mapsto x \quad \text{on } {}^*K.$$

**Claim 2.** $B^{\mu} \in A^{\mu} \in V_{k+1}(K)^{\mathbb{N}}/\sim^{\mu}$ implies $B^{\mu} \in V_k(K)^{\mathbb{N}}/\sim^{\mu}$.

*Proof of Claim 2.* $A^{\mu} \in V_{k+1}(K)^{\mathbb{N}}/\sim^{\mu}$ simply implies $A^{\mu} \in {}^*K$ or $A^{\mu} \subseteq V_k(K)^{\mathbb{N}}/\sim^{\mu}$. If $B^{\mu} \in A^{\mu}$, then $A^{\mu} \notin {}^*K$. Thus, $A^{\mu} \in V_{k+1}(K)^{\mathbb{N}}/\sim^{\mu}$ means $A^{\mu} \subseteq V_k(K)^{\mathbb{N}}/\sim^{\mu}$. Hence, $B^{\mu} \in A^{\mu} \in V_{k+1}(K)^{\mathbb{N}}/\sim^{\mu}$ implies $B^{\mu} \in V_k(K)^{\mathbb{N}}/\sim^{\mu}$. □

For every $A^{\mu} \in V_{k+1}(K)^{\mathbb{N}}/\sim^{\mu}$ and $A^{\mu} \notin {}^*K$, we set

$$j(A^{\mu}) = \left\{ j(B^{\mu}) : \left( V(K)^{\mathbb{N}}/\sim^{\mu} \models B^{\mu} \in^{\mu} A^{\mu} \right) \right\}.$$

This definition is possible if Claim 2 holds, that is, $B^{\mu} \in V_k(K)^{\mathbb{N}}/\sim^{\mu}$, which means that $j(B^{\mu})$ is defined at a previous stage of the recursive construction. This property is known as the *transitivity* of $j(V(K)^{\mathbb{N}}/\sim^{\mu})$. Hence every set in $j(V(K)^{\mathbb{N}}/\sim^{\mu})$ only consists of elements of $j(V(K)^{\mathbb{N}}/\sim^{\mu})$.
Combining $i$ and $j$;

$$ {}^*A = j(i(A)) \tag{3.10}$$

for all $A \in V(K)$.
(3.10) is the *bounded elementary embedding* of the structure of $V(K)$ into $V({}^*K)$. The membership relation $\in^{\mu}$ in the bounded ultrapower is mapped by $j$ into the ordinary membership relation in $V({}^*K)$: For all $A^{\mu}, B^{\mu} \in V(K)^{\mathbb{N}}/\sim^{\mu}$, $B^{\mu} \in^{\mu} A^{\mu} \iff j(B^{\mu}) \in j(A^{\mu})$.
In this construction, $V(K)$ and $V({}^*K)$ are connected by the Transfer Principle (see Theorem (3.4.5)).

### 3.4.1 Some basic properties of the nonstandard enlargement

Here we define some basic terms and introduce some basic properties of the Nonstandard framework.

**Definition 3.4.1.** For every $A \in V(^*K)$ where $A$ is either an element in $^*K$ or a set that belongs to $V(^*K)$:

$(a)$ $A$ is standard if $A = {}^*B$ for some $B \in V(K)$.

$(b)$ $A$ is internal if $A \in {}^*B$ for some $B \in V(K)$.

$(c)$ $A$ is external if $A$ is not internal.

It is easy to see that every standard set is internal. Standard sets are crucial but they are not very interesting. However, the internal sets and a special type of internal set known as *hyperfinite set* are very useful for applications.
Recall: A subset $A \subseteq \mathbb{N}$ is said to be finite if

$$A \subseteq \{n \in \mathbb{N} \mid n \leq m\}$$

for some $m \in \mathbb{N}$.
An internal subset $A \subseteq {}^*\mathbb{N}$ is said to be *hyperfinite* or *$^*$-finite* if

$$A \subseteq \{n \in {}^*\mathbb{N} \mid n \leq m\}$$

for some $m \in {}^*\mathbb{N}$.

**Definition 3.4.2.** (cf. Albeverio et al. [2, Definition 3.2.1]) An internal set $A \in V(^*K)$ is said to be *hyperfinite* if there is an internal one-to-one map $g$ of some proper initial segment $\{n \in {}^*\mathbb{N} \mid n \leq m\}$ of $^*\mathbb{N}$ onto $A$, where $m \in {}^*\mathbb{N}$ is called the cardinality of $A$. i.e., $|A| = m$.

Informally, hyperfinite sets are infinite sets in the nonstandard framework with all the properties and combinatorial structure of a finite set.
The next lemma is the characterization of the internal universe (see Herzberg [51]).

**Lemma 3.4.3.** For any $x \in V(^*K)$, $x \in j(V(K)^{\mathbb{N}}/\sim^{\mu})$ if and only if there exist some $y \in V(K) \setminus K$ such that $x \in {}^*y$. Then,

$$j(V(K)^{\mathbb{N}}/\sim^{\mu}) = \bigcup_{A \in V(K)\setminus K} {}^*A = \bigcup_{m \in \mathbb{N}} {}^*V_m(K).$$

*Proof.* For any bounded sequence $A$ of $(A_n)_n$ where $A_n \in V_k(K)$ for all $n$,

$${}^*A = j(i(A)).$$

If $x \in j(V(K)^{\mathbb{N}}/\sim^{\mu})$, then there exists some $k \in \mathbb{N}$ such that

$$x \in j(i(V_k(K))) = {}^*V_k(K).$$

If $y = V_k(K)$, then $x \in j(i(y))$ and we know that $j(i(y)) = {}^*y$. Thus, $x \in {}^*y$. Conversely, suppose $x \in {}^*y$ for some $y \in V(K)$. Since ${}^* = j \circ i$, ${}^*y \in j(V(K)^{\mathbb{N}}/\sim^{\mu})$. But we have shown the transitivity of $j(V(K)^{\mathbb{N}}/\sim^{\mu})$ (see the construction of the $^*$-embedding above). Thus, $x \in {}^*y$ implies that $x \in j(V(K)^{\mathbb{N}}/\sim^{\mu})$. □

Before we discuss the properties of the internal universe, we prove the Transfer Principle which says that a sentence $\psi$ holds in $V(K)$ if and only if the $^*$-*image* of $\psi$ holds in $V(^*K)$. For convenience, in the proof the superstructure is fixed over the reals: $\mathbb{R} \subset K$.

**Definition 3.4.4.** Let $\psi$ be a sentence in $\mathcal{L}_{V(\mathbb{R})}$ that holds in $V(\mathbb{R})$ with constants $a^{(1)}, \cdots, a^{(k)} \in V(\mathbb{R})$. For every $a^{*(1)}, \cdots, a^{*(k)} \in V(^*\mathbb{R})$, ${}^*\psi$ is the $^*$-image of $\psi$ in $V(^*\mathbb{R})$.

**Theorem 3.4.5** (Transfer principle)**.** Let $\psi$ be a sentence in $\mathcal{L}_{V(\mathbb{R})}$ with bounded quantifier and suppose the constants occurring in $\psi$ are $a^{(1)}, \cdots, a^{(k)}$ (for some $a^{(1)}, \ldots, a^{(k)} \in V(\mathbb{R})$). $\psi$ holds in $V(\mathbb{R})$ if and only if ${}^*\psi(a^{*(1)}, \ldots, a^{*(k)})$ holds in $V(^*\mathbb{R})$, that is,

$$V(\mathbb{R}) \models \psi \iff V(^*\mathbb{R}) \models {}^*\psi\left(\frac{a^{*(1)}}{a^{(1)}}, \cdots, \frac{a^{*(k)}}{a^{(k)}}\right).$$

*Proof.* The Transfer Principle is just the Łoś Theorem, Łoś [78], for sentences. Thus, the proof is based on slight-modification of the proof of Łoś Theorem and is done by induction on complexity of sentences that can be reduced to measure one.

We shall comment on few cases of sentences, although other cases are abbreviations of these sentences.

**Case 1.** Assuming that the sentence $\psi$ is of the form $\psi_1 \wedge \psi_2$ where $\psi_1$ and $\psi_2$ satisfies the condition of the theorem. By using the fact that the intersection of two sets with measure one has measure one (see Definition 3.2.1-(b)), and by the induction hypothesis, we have

$$\begin{aligned}
& V({}^*\mathbb{R}) \models {}^*\psi({a^*}^{(1)}, \ldots, {a^*}^{(k)}) \\
\iff & \left({}^*\psi_1({a^*}^{(1)}, \ldots, {a^*}^{(k)}) \text{ and } {}^*\psi_2({a^*}^{(1)}, \cdots, {a^*}^{(k)})\right) \\
\iff & \begin{pmatrix} \mu\{n : V(\mathbb{R}) \models \psi_1(a^{(1)}, \ldots, a^{(k)})\} = 1 \\ \text{and } \mu\{n : V(\mathbb{R}) \models \psi_2(a^{(1)}, \ldots, a^{(k)})\} = 1 \end{pmatrix} \\
\iff & \mu\{n : V(\mathbb{R}) \models \psi_1(a^{(1)}, \ldots, a^{(k)}) \wedge \psi_2(a^{(1)}, \ldots, a^{(k)})\} = 1.
\end{aligned}$$

**Case 2.** Assume the sentence $\psi$ is of the form $\neg\psi_1$, and $\psi_1$ satisfies the condition of the theorem. By using the fact that for any given set, either the set has measure one or its complement does, but not both (see Definition 3.2.1-(c)), and by the induction hypothesis, we have

$$\begin{aligned}
& V({}^*\mathbb{R}) \models {}^*\psi({a^*}^{(1)}, \ldots, {a^*}^{(k)}) \\
\iff & V({}^*\mathbb{R}) \not\models {}^*\psi_1({a^*}^{(1)}, \ldots, {a^*}^{(k)}) \\
\iff & (1 - \mu\{n : V(\mathbb{R}) \models \psi_1(a^{(1)}, \ldots, a^{(k)})\}) = 1
\end{aligned}$$

and

$$(1 - \mu\{n : V(\mathbb{R}) \models \psi_1(a^{(1)}, \ldots, a^{(k)})\}) = \mu\{n : V(\mathbb{R}) \models \neg\psi_1(a^{(1)}, \ldots, a^{(k)})\}.$$

Thus,

$$V({}^*\mathbb{R}) \models {}^*\psi({a^*}^{(1)}, \ldots, {a^*}^{(k)}) \iff \mu\{n : V(\mathbb{R}) \models \neg\psi_1(a^{(1)}, \ldots, a^{(k)})\} = 1.$$

**Case 3.** Assuming the sentence $\psi$ is of the form $(\exists x \in a(l))\, \psi_1$ where $\psi_1$ satisfies the condition of the theorem. Firstly, we note that $V({}^*\mathbb{R}) \models {}^*\psi({a^*}^{(1)}, \ldots, {a^*}^{(k)})$ if and only if there exist an internal element $a$ such that $V(^*\mathbb{R}) \models {}^*\psi_1(a, a^{*(1)}, \ldots, a^{*(k)})$:

$$\begin{aligned} &V(^*\mathbb{R}) \models {}^*\psi_1(a, a^{*(1)}, \ldots, a^{*(k)}) \\ \iff &\mu\{n : V(\mathbb{R}) \models \psi_1(a_n, a^{(1)}, \ldots, a^{(k)})\} = 1. \end{aligned} \tag{3.11}$$

We want to show that (3.11) is equivalent to

$$\begin{aligned} &V(^*\mathbb{R}) \models {}^*\psi(a, a^{*(1)}, \ldots, a^{*(k)}) \\ \iff &\mu\{n : V(\mathbb{R}) \models (\exists x \in a(l))\, \psi_1(x \in a(l), a^{(1)}, \ldots, a^{(k)})\} = 1. \end{aligned}$$

First we show there exists an $a$ such that

$$\mu\{n : V(\mathbb{R}) \models \psi_1(a_n, a^{(1)}, \ldots, a^{(k)})\} = 1$$

implies

$$\mu\{n : V(\mathbb{R}) \models (\exists x \in a(l))\, \psi_1(x \in a(l), a^{(1)}, \ldots, a^{(k)})\} = 1.$$

To do this, it is sufficient to observe that

$$\begin{aligned} &\{n : V(\mathbb{R}) \models \psi_1(a_n, a^{(1)}, \ldots, a^{(k)})\} \\ \subseteq &\{n : V(\mathbb{R}) \models (\exists x \in a(l))\, \psi_1(x \in a(l), a^{(1)}, \ldots, a^{(k)})\}. \end{aligned}$$

We know that for any set $E$, if $\mu(E) = 1$ and $E \subseteq F \subset \mathbb{N}$, then $\mu(F) = 1$. Thus,

$$\mu\{n : V(\mathbb{R}) \models (\exists x \in a(l))\, \psi_1(x \in a(l), a^{(1)}, \ldots, a^{(k)})\} = 1.$$

Conversely, we show that

$$\mu\{n : V(\mathbb{R}) \models (\exists x \in a(l))\, \psi_1(x \in a(l), a^{(1)}, \ldots, a^{(k)})\} = 1$$

implies there exists an $a$ such that $\mu\{n : V(\mathbb{R}) \models \psi_1(a_n, a^{(1)}, \ldots, a^{(k)})\} = 1$. For each $n$ in the set $\{n : V(\mathbb{R}) \models (\exists x \in a(l))\, \psi_1(x, a^{(1)}, \ldots, a^{(k)})\}$ select some element $a_n \in V(\mathbb{R})$ that oversees this, and choose $a_n$ arbitrary otherwise. We have,

$$\begin{aligned} &\mu\{n : V(\mathbb{R}) \models \psi_1(a_n, a^{(1)}, \ldots, a^{(k)})\} \\ = &\mu\{n : V(\mathbb{R}) \models (\exists x \in a(l))\, \psi_1(x \in a(l), a^{(1)}, \ldots, a^{(k)})\} \end{aligned}$$

Hence, for every given sentence $\psi$,

$$V({}^*\mathbb{R}) \models {}^*\psi({a^*}^{(1)}, \ldots, {a^*}^{(k)}) \iff \mu\{n : V(\mathbb{R}) \models \psi(a^{(1)}, \ldots, a^{(k)})\} = 1.$$

□

**Lemma 3.4.6.** For all $k \in \mathbb{N}$, ${}^*V_k(K)$ is transitive, and thus, so is ${}^*V(K)$.

*Proof.* Obviously $K$ is transitive. By induction, one can assume that $V_k(K)$ is transitive. By definition $A \in V_{k+1}(K)$ is either $A \in V_k(K)$ or $A \subseteq V_k(K)$, and $A \in V_k(K)$ is either $A \in K$ or $A \subseteq V_k(K)$. Thus, $V_{k+1}(K)$ is also transitive. $A \in V(K)$ simply implies $A \in V_k(K)$, for some $k \in \mathbb{N}$, and can be formalized by:

$$(A \in K \vee A \subseteq V_k(K)) \tag{3.12}$$

which holds in $V(K)$. By the transitivity of $V_k(K)$, it follows that $V(K)$ is transitive. Thus, $A \subseteq V_k(K) \subset V(K)$ implies $A \subset V(K)$. Applying the Transfer Principle on (3.12),

$$(A \in {}^*K \vee A \subseteq {}^*V_k(K)) \tag{3.13}$$

holds in $V({}^*K)$. Hence, ${}^*V_k(K)$ is transitive and $V({}^*K)$ is also transitive. □

It is important to note that ${}^*V_k(K) \subseteq V_k({}^*K)$ for each $k \in \mathbb{N}$. ${}^*V_k(K)$ only contain the internal objects.

**Remark 3.4.7.** If $A$ is an internal set, for every $A' \in A$, $A'$ is internal.

*Proof.* This follows directly from the transitivity of ${}^*V(K)$. □

Figure 3.1 summarizes the relation of the standard universe to its nonstandard enlargement.

Because of the canonical nature of the *-embedding (in particular it is injective) we will often identify the class of standard objects with the class of the *-images of the standard objects. According to the ambiguous, yet commonly accepted terminology, the set of hyperreals is also a standard object in a formal sense.

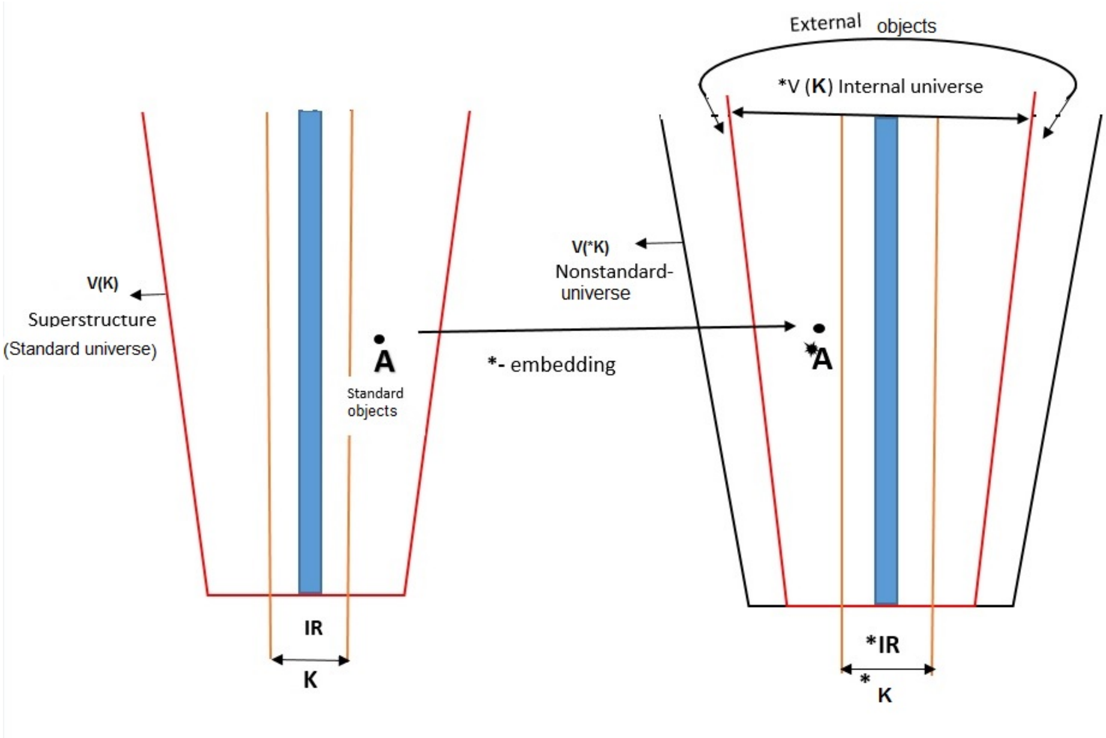


FIGURE 3.1: The standard universe and the nonstandard universe.

The next result which is the Internal Definition Principle is an important consequence of the Transfer Principle, and is the main tool for identifying an internal set.

**Theorem 3.4.8** (Internal Definition Principle)**.** Let $\psi$ be a formula with free variables $x$ and $X_1, \ldots, X_n$. Let $A_1, \ldots, A_n$ be internal sets in $V({}^*\mathbb{R})$. Then the set

$$\{y \in A_1 | \psi(y, A_1, \ldots, A_n)\} \quad \text{is internal.}$$

*Proof.* By definition, $A_1, \ldots, A_n$ are internal implies $A_1, \ldots, A_n \in {}^*V_k(\mathbb{R})$ for some $k \in \mathbb{N}$. Then, for all $X_1, \ldots X_n \in V_k(\mathbb{R})$, there is some $z \in V_{k+1}(\mathbb{R})$ such that for all $y \in V_k(\mathbb{R})$

$$y \in z \leftrightarrow y \in X_1 \wedge \psi(y, X_1, \ldots X_n) \tag{3.14}$$

holds and recall that $z = \{y \in X_1 | \psi(y, X_1, \ldots, X_n)\}$ abbreviates for all $y \in V_k(\mathbb{R})$, $y \in z \leftrightarrow y \in X_1 \wedge \psi(y, X_1, \ldots X_n)$.
Applying the Transfer Principle to (3.14), one gets that for all $X_1, \ldots X_n \in {}^*V_k(\mathbb{R})$ there is some $z \in {}^*V_{k+1}(\mathbb{R})$ such that for all $y \in {}^*V_k(\mathbb{R})$

$$y \in z \leftrightarrow y \in X_1 \wedge \psi(y, X_1, \ldots X_n). \tag{3.15}$$

By transitivity of ${}^*V_{k+1}(\mathbb{R})$ (see Lemma 3.4.6), even for all $y \in V({}^*\mathbb{R})$

$$y \in z \leftrightarrow y \in X_1 \wedge \psi(y, X_1, \ldots X_n)$$

holds. Since $A_1, \ldots, A_n \in {}^*V_k(\mathbb{R})$, we can substitute $A_j$ for $X_j$ for each $n$.
Thus, for all $y \in V({}^*\mathbb{R})$,

$$y \in z \leftrightarrow \psi(y, A_1, \ldots A_n).$$

Hence,

$$z = \{y \in A_1 | \psi(y, A_1, \ldots, A_n)\} \text{ is internal.}$$

□

The application of Transfer Principle also gives useful characterizations to many important mathematical concepts, for example, the convergence of a sequence.

**Proposition 3.4.9.** (cf. Albeverio et al. [2, Proposition 1.3.1]) Let $(a_n)_n$ be a sequence of real numbers. Then

$$\lim_{n\to\infty} a_n = a \quad \Longleftrightarrow {}^*a_k \simeq a \quad \forall\, k \in {}^*\mathbb{N} \setminus \mathbb{N}.$$

*Proof.* We assume $\lim_{n\to\infty} a_n = a$. Fix $k \in {}^*\mathbb{N} \setminus \mathbb{N}$, we want to show that $|{}^*a_k - a| < \varepsilon$, for all $\varepsilon \gg 0$. For any $\varepsilon \gg 0$, let there exists some $n \in \mathbb{N}$ such that the following holds in $V(\mathbb{R})$:

$$\forall m \in \mathbb{N}\ (m \geq n \to |a_m - a| < \varepsilon). \tag{3.16}$$

Applying the Transfer Principle on (3.16),

$$\forall m \in {}^*\mathbb{N}\ (m \geq n \to |{}^*a_m - a| < \varepsilon)$$

holds in $V({}^*\mathbb{R})$. If $k \in {}^*\mathbb{N} \setminus \mathbb{N}$, then $|{}^*a_k - a| < \varepsilon$ holds in $V({}^*\mathbb{R})$. Since this holds for all standard $\varepsilon \gg 0$, it means that ${}^*a_k \simeq a$.
Conversely, suppose ${}^*a_k \simeq a$, for all $k \in {}^*\mathbb{N} \setminus \mathbb{N}$. We fix $\varepsilon$ in $\mathbb{R}$. The set

$$S = \{n \in {}^*\mathbb{N} \mid |{}^*a_m - a| < \varepsilon, \forall m \geq n, m \in {}^*\mathbb{N}\}$$

is internal by the Internal Definition Principle and contains all $k \in {}^*\mathbb{N} \setminus \mathbb{N}$. By Underspill (cf. Albeverio et al. [2, Proposition 1.2.7]), $S$ must contain some finite $n_\varepsilon \in \mathbb{N}$. Thus, the convergence of the sequence $(a_n)_n$ in the standard sense. $\square$

Another important property of the nonstandard universe is the Countable Saturation Principle.

**Theorem 3.4.10** (Countable Saturated Principle)**.** Let $(A^i)_{i\in\mathbb{N}}$ be a decreasing sequence of nonempty internal sets such that $\bigcap\limits_{i\leq I} A^i \neq \emptyset$ for all $I \in \mathbb{N}$. Then,

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

*Proof.* Each $A^i$ is internal simply implies $A^i = j((A^i_n)^\mu)$ where $(A^i_n)^\mu$ denote the equivalence class of the sequence $(A^i_n)_n$ (see the construction of the bounded ultrapower and Lemma 3.4.3). Obviously, $A^i \neq \emptyset$, thus, one may assume that each $A^i_n \neq \emptyset$ for all $n$.

**Claim 3.** $j\Big(\bigcap\limits_{i\leq I} A^i_n\Big)^\mu = \bigcap\limits_{i\leq I} j\Big((A^i_n)^\mu\Big) = \bigcap\limits_{i\leq I} A^i.$

*Proof of Claim 3.* Let

$$A^1_n \cap A^2_n \cap \cdots \cap A^m_n = B_n \quad \text{for } i = 1, \ldots, m,$$

where $A^1_n, A^2_n, \ldots, A^m_n, B_n \in V(K)$. We want to show that

$${}^*A^1_n \cap {}^*A^2_n \cap \cdots \cap {}^*A^m_n = {}^*B_n.$$

By definition of $V(K)$, $A^1_n, A^2_n, \ldots, A^m_n, B_n \in V_k(K)$ for some $k \in \mathbb{N}$ and by transitivity of $V_k(K)$, $A^1_n, A^2_n, \ldots, A^m_n, B_n \subseteq V_k(K)$.
The expression

$$A^1_n \cap A^2_n \cap \cdots \cap A^m_n = B_n$$

can be formalized by:

$$\forall x \in V_k(K)\Big(x \in A^1_n \cap x \in A^2_n \cap \cdots \cap x \in A^m_n = x \in B_n\Big) \tag{3.17}$$

holds in $V(K)$. Applying Transfer Principle to (3.17), it becomes

$$\forall x \in {}^*V_k(K)\Big(x \in {}^*A_n^1 \cap x \in {}^*A_n^2 \cap \cdots \cap x \in {}^*A_n^m = x \in {}^*B_n\Big) \tag{3.18}$$

holds in $V({}^*K)$. Thus,

$${}^*A_n^1 \cap {}^*A_n^2 \cap \cdots \cap {}^*A_n^m = {}^*B_n. \tag{3.19}$$

From the construction of nonstandard enlargement and Lemma 3.4.3, (3.19) can be written as

$$j((A_n^1)^\mu) \cap j((A_n^2)^\mu) \cap \cdots \cap j((A_n^m)^\mu) = j((B_n)^\mu).$$

Thus,

$$j((A_n^1)^\mu) \cap j((A_n^2)^\mu) \cap \cdots \cap j((A_n^m)^\mu) = j\big((A_n^1 \cap A_n^2 \cap \cdots \cap A_n^m)^\mu\big).$$

□

From Claim 3 and by the assumption $\bigcap\limits_{i \le I} A^i \neq \emptyset$,

$$\mu\left\{n : \bigcap_{i \le I} A_n^i \neq \emptyset\right\} = 1 \quad \text{for all } I \in \mathbb{N}. \tag{3.20}$$

For each $n$, let

$$I' = \max\left\{I \in N : \bigcap_{i \le I} A_n^i \neq \emptyset \text{ and } I' \le n\right\};$$

since $A_n^1 \neq \emptyset$, $I'$ exists. Assuming an element $x_n \in \bigcap\limits_{i \le I'} A_n^i$ for each $n$; since every element of an internal set is internal, one can say that

$$j\big((x_n)^\mu\big) \in A^I$$

for all $I$, and this follows from Claim 3 and (3.20) since

$$\{n : x_n \in A_n^I\} \supset \{n : I \le I'\} = \{n : I \le n\} \bigcap \left\{n : \bigcap_{i \le I} A_n^i \neq \emptyset\right\},$$

where $\{n : I \leq n\}$ has the finite complement and thus measure one. $\square$

In the literature, a nonstandard universe that satisfies the Countable Saturation Principle is also called $\aleph_1$-*saturated.*
A quite useful consequence of the Countable Saturation Principle is the following: Recall, an internal sequence $(A_n)_{n\in{}^*\mathbb{N}}$ is the canonical extension of a sequence $(A_n)_{n\in\mathbb{N}}$ by Transfer Principle. The question is what happens if we have a countable sequence $(A_n)_{n\in\mathbb{N}}$ in $V({}^*K)$. Can we extend this sequence to an internal sequence $(A_n)_{n\in{}^*\mathbb{N}}$? The next proposition proves the useful *extension principle.*

**Proposition 3.4.11.** (cf. Albeverio et al. [2, Proposition 2.1.3]) Let $(A_n)_{n\in\mathbb{N}}$ be some bounded countable sequence of internal sets in $V({}^*K)$. Then $(A_n)_{n\in\mathbb{N}}$ can be extended to an internal sequence $(A_n)_{n\in{}^*\mathbb{N}}$ in $V({}^*K)$.

*Proof.* The sequence $(A_n)_{n\in{}^*\mathbb{N}}$ is internal simply implies there exists an internal function

$$A : {}^*\mathbb{N} \to V({}^*K)$$

such that $A(n) = A_n$ for all $n \in {}^*\mathbb{N}$. Thus, the domain of $A$ is ${}^*\mathbb{N}$ (internal) since the domain of an internal function is always internal. But the sequence $(A_n)_{n\in\mathbb{N}}$ is external, even though every element $A_n$ of the sequence is internal because $\mathbb{N}$ is external. Thus, Transfer Principle is of no use, but the countable saturation principle does the work: Let $N \in \mathbb{N}$ be such that $A_n \in {}^*V_N(K)$ for all $n \in \mathbb{N}$.

$$A'_n = \bigcap_{i=1}^{n} \{f : {}^*\mathbb{N} \to {}^*V_N(K) \mid f(i) = A(i)\}$$

is internal. Using the Countable Saturation Principle, we know that $\bigcap A'_n \neq \emptyset$. Any $f$ in this infinite intersection of all $A'_n$ is an extension of the sequence $(A_n)_{n\in\mathbb{N}}$ in $V({}^*K)$. $\square$

**Claim 4.** Let $M$ be an internal family of internal sets that is closed under finite unions. Then, $M$ is closed under hyperfinite unions too.

*Proof.* Since $M$ is internal, there must be some $W \in V(K)$ such that $M \in {}^*W$.

$$\forall M \in W \left( (\forall A, B \in M \quad A \cup B \in M) \to \forall n \in \mathbb{N} \quad \forall \vec{A} \in M^n \quad \bigcup_{k=1}^{n} A_k \in M \right). \tag{3.21}$$

Applying the Transfer Principle on (3.21),

$$\forall M \in {}^*W \left( (\forall A, B \in M \quad A \cup B \in M) \to \forall n \in {}^*\mathbb{N} \quad \forall \vec{A} \in M^n \quad \bigcup_{k=1}^{n} A_k \in M \right).$$

Thus, $M$ is closed under hyperfinite unions. $\square$

**Proposition 3.4.12.** (cf. Loeb [76, Proposition 1]) Let $(A_n)_n$ be a bounded sequence of internal sets. If $A_0 \subset \bigcup_{n=1}^{\infty} A_n$, then there exists an $m \in \mathbb{N}$ such that $A_0 \subset \bigcup_{n=1}^{m} A_n$.

*Proof.* Let $(A_n)_{n \in {}^*\mathbb{N}}$ be an internal sequence extending $(A_n)_{n \in \mathbb{N}}$ (see Proposition 3.4.11). Then,

$$S = \left\{ m \in {}^*\mathbb{N} \mid A_0 \subset \bigcup_{n=1}^{m} A_n \right\}$$

is internal and contains ${}^*\mathbb{N} \setminus \mathbb{N}$. This implies there exists some $m \in \mathbb{N}$ such that $m \in S$. Otherwise, for all $m \in \mathbb{N}$, $m \notin S$ and since $S$ is internal, by Internal Definition Principle, ${}^*\mathbb{N} \setminus \mathbb{N}$ is internal. Thus, $\mathbb{N} = {}^*\mathbb{N} \setminus ({}^*\mathbb{N} \setminus \mathbb{N})$ is internal, which is a contradiction. $\square$

However, a family of internal sets is generically not closed under countable unions. The fact that a family of internal sets is only closed under finite set operations and not under countable infinite set operations prevents immediate application of nonstandard methods to measure theory and probability theory (a $\sigma$-algebra on a set $\Omega$ is a collection of subsets of $\Omega$ that is closed under countably many set operations). But an application of the *Caratheodory extension theorem* (see next section for discussion) gives a definitive solution.

## 3.5 Nonstandard measure space to standard measure space

The basic techniques for the conversion of a nonstandard measure space to a standard measure space is the Loeb construction that mainly depends on the application of the Caratheodory extension theorem: This takes the standard part of the *internal finitely additive measure*, that is the *finitely additive measure*, and convert it into a real-valued *$\sigma$-additive measure*. We shall do this by establishing that a finitely additive measure (Definition 3.5.4) can be seen as a *premeasure* (Definition 3.5.3) and then extend the premeasure into a measure on the $\sigma$-algebra as noted by Bauer [9, Theorem 5.1]. We conclude by proving the uniqueness of this measure.

Let $(\mathbf{\Omega}, \mathcal{A}, \nu)$ be a *hyperfinite probability space*. $\mathbf{\Omega}$ is an internal set in some superstructure $V(^*K)$. $\mathcal{A}$ is an internal algebra on $\mathbf{\Omega}$, i.e., $\mathcal{A}$ is an internal set of subsets of $\mathbf{\Omega}$ which contains $\emptyset$ and $\mathbf{\Omega}$, and for every set $A, B \in \mathcal{A}$,

$$A \cup B \in \mathcal{A},$$

thus, $\mathcal{A}$ is closed under finite unions. $\mathcal{A}$ being internal implies $\mathcal{A}$ is also closed under hyperfinite unions. $\mathcal{A}$ is closed under complements, i.e.,

$$\mathbf{\Omega} \setminus A \in \mathcal{A}, \qquad \text{for every } A \in \mathcal{A}.$$

Let $\nu$ be an internal probability measure that is defined on $\mathcal{A}$. We denote by $^\circ\nu$ the standard part of $\nu$ such that $^\circ\nu$ is a finitely additive measure and takes its values in $\mathbb{R}$. Let $\sigma(\mathcal{A})$ be the smallest collection of subsets of $\mathbf{\Omega}$, both internal and external, i.e., is the $\sigma$-algebra in the standard sense containing $\mathcal{A}$. A natural question would be: under what condition does there exist a $\sigma$-algebra $\mathcal{B}$ in $\mathbf{\Omega}$ and a measure $\tilde{\mu}$ on $\mathcal{B}$ such that $^\circ\nu$ is the restriction of $\tilde{\mu}$ to $\mathcal{A}$? An obvious necessary condition would be for $^\circ\nu$ to be a premeasure on $\mathcal{A}$. A finite additive measure $^o\nu$ on the internal algebra $\mathcal{A}$ can be seen as a premeasure if the continuity property holds:

$$^o\nu(A_n) \to 0, \text{ as } n \to \infty \text{ if } A_n \downarrow \emptyset \text{ as } n \to \infty,$$

for any sequence $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{A}$. This property is trivially satisfied due to the Countable Saturated Principle: if $A_n \downarrow \emptyset$, then there exists some $k \in \mathbb{N}$ such that $A_k = \emptyset$ for all $k \geq n$. The extension theorem (Theorem 3.5.8) shows that for every premeasure $^\circ\nu$ on an internal algebra $\mathcal{A}$, there exists a $\sigma$-algebra $\mathcal{B}$ in $\mathbf{\Omega}$ with $\mathcal{A} \subset \mathcal{B}$, and a measure $\tilde{\mu}$ on $\mathcal{B}$ such that $^\circ\nu$ is the restriction of $\tilde{\mu}$ to $\mathcal{A}$. The completion of this measure is called the *Loeb measure* denoted by $L(\nu)$. The $\sigma$-algebra $\sigma(\mathcal{A})$ can be seen as the Borel-algebra of $\mathcal{A}$ and its completion $L(\mathcal{A})$ with respect to $L(\nu)$ is known as the *Loeb-algebra* of $\mathcal{A}$. Thus, the probability space $(\boldsymbol{\Omega}, L(\mathcal{A}), L(\nu))$ is called the Loeb space of $(\boldsymbol{\Omega}, \mathcal{A}, \nu)$. It suffices to say $\mathcal{B}$ is the $\sigma$-algebra $\sigma(\mathcal{A})$ generated in $\boldsymbol{\Omega}$ by an internal algebra $\mathcal{A}$.

**Definition 3.5.1.** Let $\mathcal{A}$ be a collection of subsets of some set $\boldsymbol{\Omega}$ which contains $\emptyset$ and $\boldsymbol{\Omega}$. $\mathcal{A}$ is said to be an *algebra* if $\mathcal{A}$ is closed under complement and finite union:

(a) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$;

(b) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

**Definition 3.5.2.** Let $\mathcal{A}$ be a collection of subsets of some set $\boldsymbol{\Omega}$. $\mathcal{A}$ is said to be a *$\sigma$-algebra* in $\boldsymbol{\Omega}$ if it satisfies the following properties:

(a) $\boldsymbol{\Omega} \in \mathcal{A}$;

(b) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$;

(c) $(A_n)_{n\in\mathbb{N}} \subset \mathcal{A} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

It is easy to see that $\mathcal{P}(\boldsymbol{\Omega})$ (where $\mathcal{P}(\cdot)$ denotes a power set) is always a $\sigma$-algebra.

**Definition 3.5.3.** Let $\mathcal{A}$ be an algebra. $\mu : \mathcal{A} \to \mathbb{R}_{>0} \cup \{+\infty\}$ is called a *premeasure* if

$$\mu(\emptyset) = 0 \tag{3.22}$$

and for every disjoint countable sequence $(A_n)_n$ of elements from $\mathcal{A}$ whose union lies in $\mathcal{A}$

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) \qquad \text{holds.} \tag{3.23}$$

**Definition 3.5.4.** $\mu : \mathcal{A} \to \mathbb{R}_{>0} \cup \{+\infty\}$ is said to be a *finitely additive measure* if

$$\mu\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mu(A_i) \qquad \text{holds.} \tag{3.24}$$

**Definition 3.5.5.** Let $\mathcal{A}$ be an internal algebra. $\mu : \mathcal{A} \to {}^*[0,1]$ is an *internal probability measure* if and only if

$$\mu(\Omega) = 1;$$

$$\mu(\emptyset) = 0;$$

and for all disjoint $A, B \in \mathcal{A}$,

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

**Lemma 3.5.6.** If $\nu$ is an internal probability measure, ${}^\circ\nu$ is a premeasure.

*Proof.* If $(A_n)_n$ is a disjoint countable sequence of internal sets from an internal algebra $\mathcal{A}$ such that $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$, then we have to prove that

$${}^\circ\nu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} {}^\circ\nu(A_n).$$

**Claim 5.** $\bigcup_{n\in\mathbb{N}} A_n$ is internal if and only if it equals $\bigcup_{n\leq k} A_n$ for some $k \in \mathbb{N}$.

*Proof of Claim 5.* Assuming $A = \bigcup A_n$ is internal. Consider the sequence $(B_k)_k \in \mathcal{A}$ defined by

$$B_k = A \setminus \bigcup_{n\leq k} A_n \quad \forall k \in \mathbb{N}.$$

By De Morgan's laws,

$$\bigcap_k B_k = A \setminus \bigcup_k \bigcup_{n\leq k} A_n = A \setminus \bigcup_n A_n = \emptyset.$$

By the Countable Saturation Principle, this means that $B_k = \emptyset$ for some $k$. Thus,

$$A \setminus \bigcup_n A_n = \emptyset \quad \text{but} \quad \bigcup_{n\leq k} A_n \subseteq \bigcup_n A_n = A.$$

Therefore,

$$A = \bigcup_{n\leq k} A_n.$$

□

Hence, we infer $A_{k+1} = A_{k+2} = \cdots = \emptyset$ and from Definition 3.5.3

$${}^\circ\nu(\emptyset) = 0.$$

□

**Notation 3.5.7.** Let $\mathcal{A}$ be an algebra on $\mathbf{\Omega}$. Then $\sigma(\mathcal{A})$ is the *smallest $\sigma$-algebra* $\mathcal{C}$ such that $\mathcal{A} \subseteq \mathcal{C}$.

The reason we introduced the "premeasure" will be justified in the next theorem.

**Theorem 3.5.8.** Every premeasure $\mu$ on an algebra $\mathcal{A}$ on $\mathbf{\Omega}$ can be extended to a measure $\tilde{\mu}$ on $\sigma(\mathcal{A})$.

The measure $\tilde{\mu}$ is in fact unique if $\mu(\mathbf{\Omega}) = 1$.

**Theorem 3.5.9.** Let $\mathcal{A}$ be a collection of sets that is closed under finite intersection such that $\sigma(\mathcal{A}) = \mathcal{B}$ and $\mathbf{\Omega} \in \mathcal{A}$. Then any probability measures $\mu_1$ and $\mu_2$ on $\mathcal{B}$ which satisfy

$$\mu_1(E) = \mu_2(E) \qquad \forall E \in \mathcal{A} \tag{3.25}$$

must be identical.

**Theorem 3.5.10.** Let $\nu$ be an internal probability measure on $\mathcal{A}$. Then there exists a unique measure $L(\nu)$ on $\sigma(\mathcal{A})$ such that for all $A$ in $\mathcal{A}$,

$$L(\nu)(A) = {}^{\circ}\nu(A).$$

*Proof.* By Lemma 3.5.6, ${}^{\circ}\nu$ is a premeasure and therefore, by Theorem 3.5.8 and Theorem 3.5.9 can be uniquely extended to a measure on $\sigma(\mathcal{A})$. □

For the proof of Theorem 3.5.8 and Theorem 3.5.9, the following definitions and theorems are required.

**Definition 3.5.11.** (Bauer [9, Definition 2.1]) The collection of subsets of a set $\Omega$ is called a *Dynkin system* in $\Omega$ if it has the following properties:

(a) $\Omega \in \mathcal{D}'$;

(b) $D \in \mathcal{D}' \Rightarrow D^c \in \mathcal{D}'$;

(c) For every pairwise disjoint sequence $(D_n)_n$ of elements from $\mathcal{D}'$ given $n \in \mathbb{N}$,

$$\bigcup_{n \in \mathbb{N}} D_n \in \mathcal{D}'.$$

From $(a)$ and $(b)$, we can say that $\mathcal{D}'$ contains the empty set. i.e., $(\emptyset = \Omega^c) \in \mathcal{D}'$.

**Theorem 3.5.12.** (Bauer [9, Theorem 2.3]) A Dynkin system $\mathcal{D}'$ is a $\sigma$-algebra if it is closed under finite intersection.

*Proof.* We want to show that every Dynkin system which is closed under finite intersection is a $\sigma$-algebra. We shall confirm this with Definition 3.5.2-(c) property of a $\sigma$-algebra. For any $(A_n)_n \subset \mathcal{D}'$, we have $B_0 = A_0$, $B_1 = A_1 \setminus A_0$, $B_2 = A_2 \setminus (A_1 \cup A_2)$, $\cdots$, $B_n = A_n \setminus (A_{n-1} \cup, \cdots, \cup A_0)$. By Definition 3.5.11-($b$) and the condition of the theorem, we can easily see that $B_0, B_1, B_2$ lies in $\mathcal{D}'$.

**Claim 6.**

$$\bigcup_{l \le m} B_l = \bigcup_{l \le m} A_l$$

*Proof of Claim 6.* We already know that $B_0 = A_0$. Fix $n$ and suppose that

$$\bigcup_{l \le n} B_l = \bigcup_{l \le n} A_l.$$

We want to show that

$$\bigcup_{l \le n+1} B_l = \bigcup_{l \le n+1} A_l. \tag{3.26}$$

$$\begin{aligned}
\bigcup_{l \le n+1} B_l &= \left(\bigcup_{l \le n} B_l\right) \bigcup B_{n+1} = \left(\bigcup_{l \le n} A_l\right) \bigcup B_{n+1} \\
&= \left(\bigcup_{l \le n} A_l\right) \bigcup \left(A_{n+1} \setminus \bigcup_{l \le n} A_l\right) \\
&= \left(\bigcup_{l \le n} A_l\right) \bigcup A_{n+1} = \bigcup_{l \le n+1} A_l.
\end{aligned}$$

□

Claim 6 implies that $B_{n+1} = A_{n+1} \setminus \bigcup_{l \le n} B_l$. Now, we want to prove by induction that $(B_n)_n \subset \mathcal{D}'$. We know that $B_0 \in \mathcal{D}'$ by our construction. Suppose $B_0, B_1, \cdots, B_n \in \mathcal{D}'$ for a given $n$. We want to show that $B_{n+1} \in \mathcal{D}'$. By induction hypothesis, for all $l \le n$, $B_l \in \mathcal{D}'$. By construction, $(B_l)_{l \in \mathbb{N}}$ is a pairwise disjoint sequence. Since $\mathcal{D}'$ is a Dynkin system, $\bigcup_{l \le n} B_l \in \mathcal{D}'$ lies in the Dynkin system. Thus, $A_{n+1} \setminus \bigcup_{l \le n} B_l$ is in the Dynkin system. □

**Remark 3.5.13.** Every Dynkin system is always closed under finite intersection.

*Proof.* Let $A, B \in \mathcal{D}'$. By Definition 3.5.11-$(b)$, $A^c, B^c \in \mathcal{D}'$ and $B^c \setminus A^c \in \mathcal{D}'$. Thus, $A^c \cup (B^c \setminus A^c) \in \mathcal{D}'$. $A^c \cup (B^c \setminus A^c) \Rightarrow A^c \cup B^c \Rightarrow (A \cap B)^c \in \mathcal{D}'$. Again by Definition 3.5.11-$(b)$, $A \cap B \in \mathcal{D}'$. $\square$

Let $\Gamma$ be a collection of sets that is closed under finite intersections. We can easily observe that $\mathcal{P}(\mathbf{\Omega})$ is a Dynkin system that contains $\Gamma$ and we can easily verify that for an arbitrary given family of Dynkin systems $D_i$, $\bigcap_{i \in I} D_i$ is still a Dynkin system. Thus, the intersection of the family of all Dynkin systems containing $\Gamma$ is still a Dynkin system that contains $\Gamma$.

*Proof of Theorem 3.5.8.* Let $E \subset \mathbf{\Omega}$, and let $\mathcal{U}(E)$ be the collection of all sequences $(A_n)_n$ of sets from $\mathcal{A}$ which satisfy $E \subset \bigcup\limits_{n \in \mathbb{N}} A_n$. Define,

$$\overline{\mu}(E) = \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : A_n \in \mathcal{U}(E) \right\}.$$

Then, $\overline{\mu}$ satisfies the following properties:

(a) $\overline{\mu}(\emptyset) = 0$;

(b) $E_1, E_2 \subset \mathbf{\Omega}$, $E_1 \subset E_2 \Rightarrow \overline{\mu}(E_1) \leq \overline{\mu}(E_2)$;

(c) $(E_n)_n \subset \mathbf{\Omega} \Rightarrow \overline{\mu}\left( \bigcup\limits_{n=1}^{\infty} E_n \right) \leq \sum_{n=1}^{\infty} \overline{\mu}(E_n)$.

It is obvious that $\overline{\mu} \geq 0$. We want to show that every $A \in \mathcal{A}$ is measurable. i.e.,

$$\overline{\mu}(E) \geq \overline{\mu}(E \cap A) + \overline{\mu}(E \setminus A) \quad \forall E \subseteq \mathbf{\Omega}, \tag{3.27}$$

and also show that

$$\overline{\mu}(A) = \mu(A). \tag{3.28}$$

First, we have

$$\sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty} \mu(A_n \setminus A)$$

for every sequence $(A_n)_n$ from $\mathcal{U}(E)$ as a result of the finite additivity of $\mu$. It is easy to see that $(A_n \cap A)_n$ lies in $\mathcal{U}(E \cap A)$. i.e.,

$$\begin{aligned}(A_n)_n \in \mathcal{U}(E) &\Rightarrow E \subset \bigcup_n A_n \\ &\Rightarrow E \cap A \subset \bigcup_n (A_n \cap A) \\ &\Rightarrow (A_n \cap A)_n \in \mathcal{U}(E \cap A).\end{aligned}$$

Similarly, we can show that $(A_n \setminus A)_n$ lies in $\mathcal{U}(E \setminus A)$. By the definition of $\overline{\mu}(\cdot)$,

$$\sum_{n=1}^{\infty} \mu(A_n \cap A) \geq \overline{\mu}(E \cap A) \quad \text{and} \quad \sum_{n=1}^{\infty} \mu(A_n \setminus A) \geq \overline{\mu}(E \setminus A).$$

Thus,

$$\sum_{n=1}^{\infty} \mu(A_n) \geq \overline{\mu}(E \cap A) + \overline{\mu}(E \setminus A)$$

for every sequence $(A_n)_n$. Hence, (3.27) follows.

**Claim 7.** If $\mu$ is a premeasure on $\mathcal{A}$, then for any sets $A, A_1, A_2 \ldots \in \mathcal{A}$, $A \subset \bigcup\limits_{n=1}^{\infty} A_n \Rightarrow \mu(A) \leq \sum_{n=1}^{\infty} \mu(A_n)$.

*Proof of Claim 7.* Assuming $B_n = A \cap A_n \setminus (A_{n-1} \cup \cdots A_1)$. Then $B_n \in \mathcal{A}$ and $B_n \subset A_n$. But $A$ is the disjoint union of the sequence $(B_n)_n$, and by countable additivity $\mu(A) \leq \sum_{n=1}^{\infty} \mu(B_n) \leq \sum_{n=1}^{\infty} \mu(A_n)$. □

We remark that for a given $A \in \mathcal{A}$,

$$\overline{\mu}(E) = \overline{\mu}(E \cap A) + \overline{\mu}(E \setminus A) \quad \forall E \subseteq \mathbf{\Omega}. \tag{3.29}$$

This follows directly from (3.27) and from $(c)$.
(3.28) follows on one hand from Claim 7 i.e., $\mu(A) \leq \overline{\mu}(A)$, for $A \in \mathcal{A}$, and on the other hand by considering the sequence $A, \emptyset, \emptyset, \ldots$ from $\mathcal{U}(A)$, $\overline{\mu}(A) \leq \mu(A)$.
Hence, we can conclude that every subset $A$ of $\mathbf{\Omega}$ satisfying (3.29) is $\overline{\mu}$-measurable. Let $\mathcal{A}'$ be the collection of all subsets $A$ of $\mathbf{\Omega}$ satisfying (3.29). i.e., $\mathcal{A}'$ is the set of all $\overline{\mu}$-measurable subsets of $\mathbf{\Omega}$. Now, we want to show that $\mathcal{A}'$ is a $\sigma$-algebra. From (3.29), it is easy to see that $\mathbf{\Omega} \in \mathcal{A}'$ and whenever $A$ lies in $\mathcal{A}'$, $A^c$ also lies

in $\mathcal{A}'$: $A \in \mathcal{A}'$ implies

$$\overline{\mu}(E) = \overline{\mu}(E \cap A) + \overline{\mu}(E \setminus A) \quad \forall E \subseteq \mathbf{\Omega}$$

and $A^c \in \mathcal{A}'$ implies

$$\overline{\mu}(E) = (E \setminus A) + \overline{\mu}(E \cap A) \quad \forall E \subseteq \mathbf{\Omega}.$$

To confirm the Definition 3.5.2-(c) property of a $\sigma$-algebra, we begin by proving that the union of any two sets of $\mathcal{A}'$ also lies in $\mathcal{A}'$, and so $\mathcal{A}'$ is an algebra. $B \in \mathcal{A}'$ implies

$$\overline{\mu}(E) = \overline{\mu}(E \cap B) + \overline{\mu}(E \setminus B) \quad \forall E \subseteq \mathbf{\Omega}. \tag{3.30}$$

Then we split (3.30) into two different equations. i.e.,

$$\overline{\mu}(E \cap B) = \overline{\mu}(E \cap B \cap A) + \overline{\mu}(E \cap B \cap A^c) \quad \forall E \subseteq \mathbf{\Omega}.$$

and

$$\overline{\mu}(E \setminus B) = \overline{\mu}(E \cap B^c \cap A) + \overline{\mu}(E \cap B^c \cap A^c) \quad \forall E \subseteq \mathbf{\Omega}.$$

Thus,

$$\overline{\mu}(E) = \overline{\mu}(E \cap B \cap A) + \overline{\mu}(E \cap B \cap A^c) + \overline{\mu}(E \cap B^c \cap A) + \overline{\mu}(E \cap B^c \cap A^c). \tag{3.31}$$

Replacing $E$ by $E \cap (A \cup B)$ in (3.31), we have

$$\overline{\mu}(E \cap (A \cup B)) = \overline{\mu}(E \cap B \cap A) + \overline{\mu}(E \cap B \cap A^c) + \overline{\mu}(E \cap A \cap B^c) \tag{3.32}$$

Substituting (3.32) into (3.31),

$$\overline{\mu}(E) = \overline{\mu}(E \cap (A \cup B)) + \overline{\mu}(E \cap A^c \cap B^c).$$

Thus, $A \cup B \in \mathcal{A}'$ for all $E \subseteq \mathbf{\Omega}$.

Now, let $(A_n)_n$ be a sequence of pairwise disjoint sets from $\mathcal{A}'$ and let $A = \bigcup_{n \in \mathbb{N}} A_n$. Put $A = A_1$ and $B = A_2$ in (3.32), then (3.32) becomes

$$\overline{\mu}(E \cap (A_1 \cup A_2)) = \overline{\mu}(E \cap A_1 \cap A_2^c) + \overline{\mu}(E \cap A_1^c \cap A_2) \quad \forall E \subseteq \mathbf{\Omega}$$

and

$$\overline{\mu}(E \cap (A_1 \cup A_2)) = \overline{\mu}(E \cap A_1) + \overline{\mu}(E \cap A_2). \tag{3.33}$$

By induction, we can generalize (3.33):

$$\overline{\mu}(E \cap \bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} (E \cap A_i)$$

for all $E \subseteq \mathbf{\Omega}$ and for all $n$. Recall: we have already proved that $\mathcal{A}'$ is closed under finite union. i.e., $B_n = \bigcup A_i$ lies in $\mathcal{A}'$. $E \backslash B_n \supset E \backslash A$, so that $\overline{\mu}(E \backslash B_n) \geq \overline{\mu}(E \backslash A)$, we obtain

$$\overline{\mu}(E) = \overline{\mu}(E \cap B_n) + \overline{\mu}(E \setminus B_n) \geq \sum_{i=1}^{n} \overline{\mu}(E \cap A_i) + \overline{\mu}(E \setminus A) \tag{3.34}$$

for all $n$. From (3.34) and using $(c)$,

$$\overline{\mu}(E) \geq \sum_{i=1}^{n} \overline{\mu}(E \cap A_n) + \overline{\mu}(E \setminus A) \geq \overline{\mu}(E \cap A) + \overline{\mu}(E \setminus A)$$

holds for all $E \subseteq \mathbf{\Omega}$ and then we have (using the equality condition as in (3.29))

$$\overline{\mu}(E) = \sum_{i=1}^{n} \overline{\mu}(E \cap A_n) + \overline{\mu}(E \setminus A) = \overline{\mu}(E \cap A) + \overline{\mu}(E \setminus A) \tag{3.35}$$

for all $E \subseteq \mathbf{\Omega}$. Thus, $A = \bigcup_{n \in \mathbb{N}} A_n$ lies in $\mathcal{A}'$. Hence, we can say that the algebra $\mathcal{A}'$ is a Dynkin system that is closed under intersection. By Theorem 3.5.12, $\mathcal{A}'$ is a $\sigma$-algebra. If we set $E = A$ in (3.35),

$$\overline{\mu}(A) = \sum_{n=1}^{\infty} \overline{\mu}(A_n).$$

We conclude that the restriction of $\overline{\mu}$ to $\mathcal{A}'$ is a measure. $\square$

The summary of what we have shown is that every subset $A$ of $\mathbf{\Omega}$ satisfying (3.29) is $\overline{\mu}$-measurable. $\mathcal{A}'$ is a $\sigma$-algebra. (3.29) has shown that $\mathcal{A} \subset \mathcal{A}'$, thus, $\sigma(\mathcal{A}) \subset \mathcal{A}'$. (3.28) implies that $\tilde{\mu} := \overline{\mu}$ given $\sigma(\mathcal{A})$ is an extension of $\mu$ to a measure on $\sigma(\mathcal{A})$.

**Definition 3.5.14.** Let $\Gamma$ be a collection of set that is closed under finite intersection. Every $\Gamma \subset \mathcal{P}(\mathbf{\Omega})$ lies in a smallest Dynkin system. This smallest Dynkin system is called the Dynkin system generated by $\Gamma$, and it is denoted by $\delta(\Gamma)$.

**Theorem 3.5.15.** Every $\mathcal{A} \subset \mathcal{P}(\mathbf{\Omega})$ which is closed under finite intersection satisfies

$$\delta(\mathcal{A}) = \sigma(\mathcal{A}).$$

*Proof.* By definition, every $\sigma$-algebra is a Dynkin system. Thus, $\sigma(\mathcal{A})$ is a Dynkin system that contain $\mathcal{A}$. On the one hand, $\delta(\mathcal{A}) \subset \sigma(\mathcal{A})$, since $\delta(\mathcal{A})$ is the smallest Dynkin system containing $\mathcal{A}$. On the other hand, $\sigma(\mathcal{A}) \subset \delta(\mathcal{A})$. This follows directly from Remark 3.5.13. □

*Proof of Theorem 3.5.9.* Let $E \in \mathcal{A}$ such that $\mu_1(E) = \mu_2(E)$. Consider the set

$$\mathcal{D}_E := \{D \in \mathcal{B} : \mu_1(E \cap D) = \mu_2(E \cap D)\}. \tag{3.36}$$

We want to show that $\mathcal{D}_E$ is a Dynkin system. Obviously, $\mathbf{\Omega} \in \mathcal{D}_E$. If $D \in \mathcal{D}_E$, then

$$\mu_1(E \cap D) = \mu_2(E \cap D).$$

We know that,

$$\begin{aligned}\mu_1(E \cap D^c) &= \mu_1(E \setminus (E \cap D)) = \mu_1(E) - \mu_1(E \cap D)\\ &= \mu_2(E) - \mu_2(E \cap D) = \mu_2(E \cap D^c)\end{aligned}$$

which implies that $D^c \in \mathcal{D}_E$. This satisfies Definition 3.5.11-(b). Definition 3.5.11-(c) follows from the $\sigma$-additivity of measures $\mu_1$ and $\mu_2$. Since $\mathcal{A}$ is closed under finite intersections, $\mathcal{A} \subset \mathcal{D}_E$ follows from (3.25) and (3.36). But then $\delta(\mathcal{A}) \subset \mathcal{D}_E$ since $\delta(\mathcal{A})$ is the smallest Dynkin system generated by $\mathcal{A}$. By Theorem 3.5.15, we know that every $\mathcal{A} \subset \mathcal{P}(\mathbf{\Omega})$ which is closed under finite intersection satisfies $\delta(\mathcal{A}) = \sigma(\mathcal{A})$. This implies $\delta(\mathcal{A}) = \sigma(\mathcal{A}) = \mathcal{B}$. Therefore, $\delta(\mathcal{A}) \subset \mathcal{D}_E \subset \mathcal{B}$ implies $\mathcal{D}_E = \mathcal{B}$ for all $E \in \mathcal{A}$ satisfying $\mu_1(E) = \mu_2(E)$, in particular for $E = \mathbf{\Omega}$. □

# Chapter 4

# Hyperfinite Construction of $G$-expectation

## 4.1 Introduction

The hyperfinite $G$-expectation is a nonstandard discrete analogue of $G$-expectation (in the sense of Robinsonian nonstandard analysis) which is infinitely close to the continuous time $G$-expectation. We develop the basic theory for the hyperfinite $G$-expectation. We prove a lifting theorem for the $G$-expectation. Herein, we use an existing discretization theorem for the $G$-expectation from Chapter 2, Theorem 2.3.13. Very roughly speaking, we extend the discrete time analogue of $G$-expectation to a hyperfinite time analogue. Then, we use the characterization of convergence in nonstandard analysis to prove that the hyperfinite discrete-time analogue of the $G$-expectation is infinitely close to the standard $G$-expectation.

Nonstandard analysis makes consistent use of infinitesimals in mathematical analysis based on techniques from mathematical logic. This approach is very promising because it also allows, for instance, to study continuous-time stochastic processes as formally finite objects. Many authors have applied nonstandard analysis to problems in measure theory, probability theory and mathematical economics (see for example, Anderson and Raimondo [5] and the references therein or the contribution in Berg [12]), especially after Loeb [76] converted nonstandard measures (i.e. the images of standard measures under the nonstandard embedding $^*$) into real-valued, countably additive measures, by means of the standard part operator and *Caratheodory*'s extension theorem. One of the main ideas behind these applications is the extension of the notion of a finite set known as *hyperfinite set* or more causally, a formally finite set. Very roughly speaking, hyperfinite sets are sets that can be formally enumerated with both standard and nonstandard natural numbers up to a (standard or nonstandard, i.e. unlimited) natural number.

Anderson [3], Keisler [58], Lindstrøm [70], Hoover and Perkins [54], a few to mention, used Loeb's [76] approach to develop basic nonstandard stochastic analysis and in particular, the nonstandard Itô calculus. Loeb [76] also presents the construction of a Poisson processes using nonstandard analysis. Anderson [3] showed that Brownian motion can be constructed from a hyperfinite number of coin tosses, and provides a detailed proof using a special case of Donsker's theorem. Anderson [3] also gave a nonstandard construction of stochastic integration with respect to his construction of Brownian motion. Keisler [58] uses Anderson's [3] result to obtain some results on stochastic differential equations. Lindstrøm [73] gave the hyperfinite construction (*lifting*) of $L^2$ standard martingales. Using nonstandard stochastic analysis, Perkins [90] proved a global characterization of (standard) Brownian local time. In this chapter, we do not work on the Loeb space because the $G$-expectation and its corresponding $G$-Brownian motion is not based on a classical probability measure, but on a set of martingale laws.

Dolinsky et al. [38] and Chapter 2 (Theorem 2.3.13) showed the standard weak approximation of the $G$-expectation. Dolinsky et al. [38] introduced a notion of volatility uncertainty in discrete time and defined a discrete version of Peng's $G$-expectation. In the continuous-time limit, it turns out that the resulting sublinear expectation converges weakly to the $G$-expectation. To allow for the hyperfinite construction of $G$-expectation which require a discretization of the state space, in Chapter 2 we refine the discretization by Dolinsky et al. [38] and obtain a discretization where the martingale laws are defined on a finite lattice rather than the whole set of reals.

The aim of this chapter is to give an alternative, combinatorially inspired construction of the $G$-expectation based on the aforementioned Theorem 2.3.13. We hope that this result may eventually become useful for applications in financial economics (especially existence of equilibrium on continuous-time financial markets with volatility uncertainty) and provides additional intuition for *Shige Peng*'s $G$-stochastic analysis. We begin the nonstandard treatment of the $G$-expectation by defining a notion of $S$-continuity, a standard part operator, and proving a corresponding lifting (and pushing down) theorem. Thereby, we show that our hyperfinite construction is the appropriate nonstandard analogue of the $G$-expectation. For details on nonstandard analysis, we refer the reader to Cutland [28], Albeverio et al. [2], Loeb and Wolff [77] and Stroyan and Luxemburg [103].

The rest of this chapter is organised as follows: in Section 4.2, we introduce the $G$-expectation, the continuous-time setting of the sublinear expectation and the hyperfinite-time setting needed for our construction. In Section 4.3, we introduce the notion of $S$-continuity and also define the appropriate lifting notion needed for our construction. Finally, we prove that the hyperfinite $G$-expectation is infinitely close to the standard $G$-expectation.

## 4.2 Framework

The $G$-expectation $\xi \mapsto \mathcal{E}^G(\xi)$ is a sublinear function that takes random variables on the canonical space $\Omega$ to the real numbers. The symbol $G$ is a function $G : \mathbb{R} \to \mathbb{R}$ of the form

$$G(\gamma) = \frac{1}{2} \sup_{c \in \mathbf{D}} c\gamma, \tag{4.1}$$

where $\mathbf{D} = [r_{\mathbf{D}}, R_{\mathbf{D}}]$ and $0 \leq r_{\mathbf{D}} \leq R_{\mathbf{D}} < \infty$. Let $\mathcal{P}^G$ be the set of probabilities on $\Omega$ such that for any $P \in \mathcal{P}^G$, $B$ is a martingale with volatility $d\langle B\rangle_t / dt \in \mathbf{D}$ in $P \otimes dt$ a.e. Then, the dual view of the $G$-expectation via volatility uncertainty (cf. Denis et al. [37]) can be denoted as

$$\mathcal{E}^G(\xi) = \sup_{P \in \mathcal{P}^G} \mathbb{E}^P[\xi].$$

The canonical process $B$ under the $G$-expectation $\mathcal{E}^G$ is called *$G$-Brownian motion* (cf. Peng [89]).

### 4.2.1 Continuous-time construction of sublinear expectation

Let $\Omega = \{\omega \in \mathcal{C}([0,T]; \mathbb{R}) : \omega_0 = 0\}$ be the canonical space of continuous paths on $[0,T]$ endowed with the maximum norm $\|\omega\|_\infty = \sup_{0\leq t\leq T} |\omega_t|$, where $|\cdot|$ is the Euclidean norm on $\mathbb{R}$. $B$ is the canonical process defined by $B_t(\omega) = \omega_t$ and $\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t)$ is the filtration generated by $B$. $\mathcal{P}_{\mathbf{D}}$ is the set of all martingale laws on $\Omega$ such that under any $P \in \mathcal{P}_{\mathbf{D}}$, the coordinate process $B$ is a martingale with respect to $\mathcal{F}_t$ with volatility $d\langle B\rangle_t/dt$ taking values in $\mathbf{D}$, $P \otimes dt$ a.e., for $\mathbf{D} = [r_{\mathbf{D}}, R_{\mathbf{D}}]$ and $0 \leq r_{\mathbf{D}} \leq R_{\mathbf{D}} < \infty$.

$$\mathcal{P}_{\mathbf{D}} = \{P \text{ martingale law on } \Omega;\ d\langle B\rangle_t/dt \in \mathbf{D},\ P \otimes dt \text{ a.e.}\}.$$

Thus, the sublinear expectation is given by

$$\mathcal{E}_{\mathbf{D}}(\xi) = \sup_{P \in \mathcal{P}_{\mathbf{D}}} \mathbb{E}^P[\xi], \tag{4.2}$$

for any $\xi : \Omega \to \mathbb{R}$, $\xi$ is $\mathcal{F}_T$-measurable and integrable for all $P \in \mathcal{P}_{\mathbf{D}}$. Here $\mathbb{E}^P$ denotes the expectation under $P$. It is important to note that the continuous-time sublinear expectation (4.2) coincides with the classical $G$-expectation (for every $\xi \in \mathbb{L}^1_G$ where $\mathbb{L}^1_G$ is defined as the $\mathbb{E}[|\cdot|]-$norm completion of $\mathcal{C}_b(\Omega;\mathbb{R})$) provided (4.1) is satisfied see Chapter 2.

### 4.2.2 Hyperfinite-time setting

Here we present the nonstandard version of the discrete-time setting of the sublinear expectation and the strong formulation of volatility uncertainty on the hyperfinite timeline. For the standard strong formulation of volatility uncertainty in the discrete-time and continuous-time settings see Chapter 2.

**Definition 4.2.1.** ${}^*\Omega$ is the ${}^*$-image of $\Omega$ endowed with the ${}^*$-extension of the maximum norm ${}^*\|\cdot\|_\infty$.

${}^*\mathbf{D} = {}^*[r_{\mathbf{D}}, R_{\mathbf{D}}]$ is the ${}^*$-image of $\mathbf{D}$, and as such it is internal.

It is important to note that $st : {}^*\Omega \to \Omega$ is the standard part map, and $st(\omega)$ will be referred to as the *standard part* of $\omega$, for every $\omega \in {}^*\Omega$. ${}^\circ z$ denotes the standard part of a hyperreal $z$.

**Definition 4.2.2.** $\widetilde{\omega} \in {}^*\Omega$ is a *nearstandard point* if there exists $\omega \in \Omega$ such that ${}^*\|\widetilde{\omega} - {}^*\omega\|_\infty \simeq 0$. We denote the set of all nearstandard elements in ${}^*\Omega$ with $ns({}^*\Omega)$.

For all hypernatural $N$, let

$$\mathcal{L}_N = \left\{ \frac{K}{N\sqrt{N}}, \quad -N^2\sqrt{R_{\mathbf{D}}} \le K \le N^2\sqrt{R_{\mathbf{D}}}, \quad K \in {}^*\mathbb{Z} \right\} \tag{4.3}$$

and the hyperfinite timelime

$$\mathbb{T} = \left\{ 0, \frac{T}{N}, \cdots, -\frac{T}{N} + T, T \right\}. \tag{4.4}$$

We consider $\mathcal{L}_N^{\mathbb{T}}$ as the canonical space of paths on the hyperfinite timeline, and $X^N = (X_k^N)_{k=0}^N$ as the canonical process denoted by $X_k^N(\bar{\omega}) = \bar{\omega}_k$ for $\bar{\omega} \in \mathcal{L}_N^{\mathbb{T}}$. $\mathcal{F}^N$ is the internal filtration generated by $X^N$. The linear interpolation operator can be written as

$$\widetilde{\ } : \widehat{\cdot} \circ \iota^{-1} \to {}^*\Omega, \quad \text{for } \widetilde{\mathcal{L}_N^{\mathbb{T}}} \subseteq {}^*\Omega,$$

where

$$\widehat{\omega}(t) := (\lfloor Nt/T \rfloor + 1 - Nt/T)\omega_{\lfloor Nt/T \rfloor} + (Nt/T - \lfloor Nt/T \rfloor)\omega_{\lfloor Nt/T \rfloor + 1},$$

for $\omega \in \mathcal{L}_N^{N+1}$ and for all $t \in {}^*[0,T]$. $\lfloor y \rfloor$ denotes the greatest integer less than or equal to $y$ and $\iota : \mathbb{T} \to \{0, \cdots, N\}$ for $\iota : t \mapsto Nt/T$.

For the hyperfinite strong formulation of the volatility uncertainty, fix $N \in {}^*\mathbb{N} \setminus \mathbb{N}$. Consider $\left\{\pm\frac{1}{\sqrt{N}}\right\}^{\mathbb{T}}$, and let $P_N$ be the uniform counting measure on $\left\{\pm\frac{1}{\sqrt{N}}\right\}^{\mathbb{T}}$. $P_N$ can also be seen as a measure on $\mathcal{L}_N^{\mathbb{T}}$, concentrated on $\left\{\pm\frac{1}{\sqrt{N}}\right\}^{\mathbb{T}}$. Let $\Omega_N = \{\underline{\omega} = (\underline{\omega}_1, \cdots, \underline{\omega}_N); \underline{\omega}_i = \{\pm 1\}, i = 1, \cdots, N\}$, and let $\Xi_1, \cdots, \Xi_N$ be a $^*$-independent sequence of $\{\pm 1\}$-valued random variables on $\Omega_N$ and the components of $\Xi_k$ are orthonormal in $L^2(P_N)$. We denote the hyperfinite random walk by

$$\mathbb{X}_t = \frac{1}{\sqrt{N}} \sum_{l=1}^{Nt/T} \Xi_l \quad \text{for all } t \in \mathbb{T}.$$

The hyperfinite-time stochastic integral of some $F : \mathbb{T} \times \mathcal{L}_N^{\mathbb{T}} \to {}^*\mathbb{R}$ with respect to the hyperfinite random walk is given by

$$\sum_{s=0}^{t} F(s, \mathbb{X})\Delta\mathbb{X}_s : \Omega_N \to {}^*\mathbb{R}, \quad \underline{\omega} \in \Omega_N \mapsto \sum_{s=0}^{t} F(s, \mathbb{X}(\underline{\omega}))\Delta\mathbb{X}_s(\underline{\omega}).$$

Thus, the hyperfinite set of martingale laws can be defined by

$$\bar{\mathcal{Q}}^N_{\mathbf{D}'_N} = \left\{ \; P_N \circ (M^{F,\mathbb{X}})^{-1}; \; F : \mathbb{T} \times \mathcal{L}^{\mathbb{T}}_N \to \sqrt{\mathbf{D}'_N} \; \right\}$$

where

$$\mathbf{D}'_N = {}^*\mathbf{D} \cap \left( \frac{1}{N} {}^*\mathbb{N} \right)^2$$

and

$$M^{F,\mathbb{X}} = \left( \sum_{s=0}^{t} F(s, \mathbb{X}) \Delta \mathbb{X}_s \right)_{t \in \mathbb{T}} .$$

**Remark 4.2.3.** Up to scaling, $\bar{\mathcal{Q}}^N_{\mathbf{D}'_N} = \mathcal{Q}^n_{\mathbf{D}'_n}$.

## 4.3 Results and proofs

**Definition 4.3.1** (Uniform lifting of $\xi$)**.** Let $\Xi : \mathcal{L}^{\mathbb{T}}_N \to {}^*\mathbb{R}$ be an internal function, and let $\xi : \Omega \to \mathbb{R}$ be a continuous function. $\Xi$ is said to be a *uniform lifting* of $\xi$ if and only if

$$\forall \bar{\omega} \in \mathcal{L}^{\mathbb{T}}_N \Big( \widetilde{\bar{\omega}} \in ns({}^*\Omega) \Rightarrow {}^\circ\Xi(\bar{\omega}) = \xi(st(\widetilde{\bar{\omega}})) \Big),$$

where $st(\widetilde{\bar{\omega}})$ is defined with respect to the topology of uniform convergence on $\Omega$.

In order to construct the hyperfinite version of the $G$-expectation, we need to show that the $^*$-image of $\xi$, ${}^*\xi$, with respect to $\widetilde{\bar{\omega}} \in ns({}^*\Omega)$, is the canonical lifting of $\xi$ with respect to $st(\widetilde{\bar{\omega}}) \in \Omega$. i.e., for every $\widetilde{\bar{\omega}} \in ns({}^*\Omega)$, ${}^\circ \left( {}^*\xi(\widetilde{\bar{\omega}}) \right) = \xi(st(\widetilde{\bar{\omega}}))$. To do this, we need to show that ${}^*\xi$ is S-continuous in every nearstandard point $\widetilde{\bar{\omega}}$.

**Remark 4.3.2.** The following are equivalent for an internal function $\Phi : {}^*\Omega \to {}^*\mathbb{R}$:

(1) $\forall \omega' \in {}^*\Omega \left( {}^*\|\omega - \omega'\|_\infty \simeq 0 \Rightarrow {}^*|\Phi(\omega) - \Phi(\omega')| \simeq 0 \right)$.

(2) $\forall \varepsilon \gg 0, \exists \delta \gg 0 : \forall \omega' \in {}^*\Omega \left( {}^*\|\omega - \omega'\|_\infty < \delta \Rightarrow {}^*|\Phi(\omega) - \Phi(\omega')| < \varepsilon \right)$.

*Proof.* Let $\Phi$ be an internal function such that condition (1) holds. To show that (1) $\Rightarrow$ (2), fix $\varepsilon \gg 0$. We shall show there exists a $\delta$ for this $\varepsilon$ as in condition (2). Since $\Phi$ is internal, the set

$$I = \left\{ \delta \in {}^*\mathbb{R}_{>0} : \; \forall \omega' \in {}^*\Omega \; ({}^*\|\omega - \omega'\|_\infty < \delta \Rightarrow {}^*|\Phi(\omega) - \Phi(\omega')| < \varepsilon) \right\},$$

is internal by the Internal Definition Principle (see Theorem 3.4.8) and also contains every positive infinitesimal. By Underspill (cf. Albeverio et al. [2, Proposition 1.27]) $I$ must then contain some positive $\delta \in \mathbb{R}$.
Conversely, suppose condition (1) does not hold, that is, there exists some $\omega' \in {}^*\Omega$ such that

$$^*\|\omega - \omega'\|_\infty \simeq 0 \text{ and } {}^*|\Phi(\omega) - \Phi(\omega')| \text{ is not infinitesimal.}$$

If $\varepsilon = \min(1, {}^*|\Phi(\omega) - \Phi(\omega')|/2)$, we know that for each standard $\delta > 0$, there is a point $\omega'$ within $\delta$ of $\omega$ at which $\Phi(\omega')$ is farther than $\varepsilon$ from $\Phi(\omega)$. This shows that condition (2) cannot hold either. □

(The case of Remark 4.3.2 where $\Omega = \mathbb{R}$ is well known and proved in Stroyan and Luxemburg [103, Theorem 5.1.1])

**Definition 4.3.3.** Let $\Phi : {}^*\Omega \to {}^*\mathbb{R}$ be an internal function. We say $\Phi$ is *$S$-continuous* in $\omega \in {}^*\Omega$, if and only if it satisfies one of the two equivalent conditions of Remark 4.3.2.

**Proposition 4.3.4.** If $\xi : \Omega \to \mathbb{R}$ is a continuous function satisfying $|\xi(\omega)| \leq a(1 + \|\omega\|_\infty)^b$, for $a, b > 0$, then, $\Xi = {}^*\xi \circ \tilde{\cdot}$ is a uniform lifting of $\xi$.

*Proof.* Fix $\omega \in \Omega$. By definition, $\xi$ is continuous on $\Omega$. i.e., for all $\omega \in \Omega$, and for every $\varepsilon \gg 0$, there is a $\delta \gg 0$, such that for every $\omega' \in \Omega$, if

$$\|\omega - \omega'\|_\infty < \delta, \text{ then } |\xi(\omega) - \xi(\omega')| < \varepsilon. \tag{4.5}$$

By the Transfer Principle (see Theorem 3.4.5): For all $\omega \in \Omega$, and for every $\varepsilon \gg 0$, there is a $\delta \gg 0$, such that for every $\omega' \in {}^*\Omega$, (4.5) becomes,

$$^*\|{}^*\omega - \omega'\|_\infty < \delta, \text{ and } {}^*|{}^*\xi({}^*\omega) - {}^*\xi(\omega')| < \varepsilon. \tag{4.6}$$

So, ${}^*\xi$ is $S$-continuous in ${}^*\omega$ for all $\omega \in \Omega$. Applying the equivalent characterization of $S$-continuity, Remark 4.3.2, (4.6) can be written as

$$^*\|{}^*\omega - \omega'\|_\infty \simeq 0, \text{ and } {}^*|{}^*\xi({}^*\omega) - {}^*\xi(\omega')| \simeq 0.$$

We assume $\widetilde{\omega}$ to be a nearstandard point. By Definition 4.2.2, this simply implies,

$$\forall \widetilde{\omega} \in ns(^*\Omega),\ \exists \omega \in \Omega : {}^*\|\widetilde{\omega} - {}^*\omega\|_\infty \simeq 0. \tag{4.7}$$

Thus, by $S$-continuity of ${}^*\xi$ in ${}^*\omega$,

$$^*|{}^*\xi(\widetilde{\omega}) - {}^*\xi({}^*\omega)| \simeq 0.$$

Using the triangle inequality, if $\omega^{'} \in {}^*\Omega$ with ${}^*\|\widetilde{\omega} - \omega^{'}\|_\infty \simeq 0$,

$$^*\|{}^*\omega - \omega^{'}\|_\infty \leq {}^*\|{}^*\omega - \widetilde{\omega}\|_\infty + {}^*\|\widetilde{\omega} - \omega^{'}\|_\infty \simeq 0$$

and therefore again by the $S$-continuity of ${}^*\xi$ in ${}^*\omega$,

$$^*|{}^*\xi({}^*\omega) - {}^*\xi(\omega^{'})| \simeq 0.$$

And so,

$$^*|{}^*\xi(\widetilde{\omega}) - {}^*\xi(\omega^{'})| \leq {}^*|{}^*\xi(\widetilde{\omega}) - {}^*\xi({}^*\omega)| + {}^*|{}^*\xi({}^*\omega) - {}^*\xi(\omega^{'})| \simeq 0.$$

Thus, for all $\widetilde{\omega} \in ns(^*\Omega)$ and $\omega^{'} \in {}^*\Omega$, if ${}^*\|\widetilde{\omega} - \omega^{'}\|_\infty \simeq 0$, then,

$$^*|{}^*\xi(\widetilde{\omega}) - {}^*\xi(\omega^{'})| \simeq 0.$$

Hence, ${}^*\xi$ is S-continuous in $\widetilde{\omega}$. Equation (4.7) also implies

$$\widetilde{\omega} \in m(\omega) \left( m(\omega) = \bigcap \{{}^*\mathcal{O}; \mathcal{O} \text{ is an open neighbourhood of } \omega\} \right)$$

such that $\omega$ is unique, and in this case $st(\widetilde{\omega}) = \omega$.
Therefore,

$$^\circ\left({}^*\xi(\widetilde{\omega})\right) = \xi(st(\widetilde{\omega})).$$

□

**Definition 4.3.5.** Let $\bar{\mathcal{E}} : {}^*\mathbb{R}^{\mathcal{L}_N^{\mathbb{T}}} \to {}^*\mathbb{R}$. We say that $\bar{\mathcal{E}}$ lifts $\mathcal{E}^G$ if and only if for every $\xi : \Omega \to \mathbb{R}$ that satisfies $|\xi(\omega)| \leq a(1 + \|\omega\|_\infty)^b$ for some $a, b > 0$,

$$\bar{\mathcal{E}}({}^*\xi \circ \tilde{\cdot}) \simeq \mathcal{E}^G(\xi).$$

**Theorem 4.3.6.**

$$\max_{\bar{Q}\in\bar{\mathcal{Q}}^N_{\mathbf{D}'_N}} \mathbb{E}^{\bar{Q}}[\cdot] \text{ lifts } \mathcal{E}^G(\xi). \tag{4.8}$$

*Proof.* From the standard approximation in Theorem 2.3.13,

$$\max_{\mathbb{Q}\in\mathcal{Q}^n_{\mathbf{D}'_n}} \mathbb{E}^{\mathbb{Q}}[\xi(\widehat{X}^n)] \to \mathcal{E}^G(\xi), \quad \text{as } n\to\infty. \tag{4.9}$$

For all $N\in {}^*\mathbb{N}\setminus\mathbb{N}$, we know that (4.9) holds if and only if

$$\max_{Q\in{}^*\mathcal{Q}^N_{\mathbf{D}'_N}} \mathbb{E}^{Q}[{}^*\xi(\widehat{X}^N)] \simeq \mathcal{E}^G(\xi), \tag{4.10}$$

see Proposition 3.4.9. Now, we want to express (4.10) in term of $\bar{\mathcal{Q}}^N_{\mathbf{D'_N}}$. i.e., to show that

$$\max_{\bar{Q}\in\bar{\mathcal{Q}}^N_{\mathbf{D}'_N}} \mathbb{E}^{\bar{Q}}[{}^*\xi\circ\tilde{\cdot}] \simeq \mathcal{E}^G(\xi).$$

To do this, use

$$\mathbb{E}^{Q}[{}^*\xi\circ\hat{\cdot}] = \mathbb{E}^{Q}[{}^*\xi\circ\hat{\cdot}\circ\iota^{-1}\circ\iota]$$

and

$$\begin{aligned}
\mathbb{E}^{Q}[{}^*\xi\circ\hat{\cdot}\circ\iota^{-1}\circ\iota] &= \mathbb{E}^{Q}[{}^*\xi\circ\tilde{\cdot}\circ\iota] \\
&= \int_{{}^*\mathbb{R}^{N+1}} {}^*\xi\circ\tilde{\cdot}\circ\iota\, dQ, \quad \text{(transforming measure)} \\
&= \int_{{}^*\mathbb{R}^{\mathbb{T}}} {}^*\xi\circ\tilde{\cdot}\, d(Q\circ j), \\
&= \mathbb{E}^{Q\circ j}[{}^*\xi\circ\tilde{\cdot}]
\end{aligned}$$

for $j: {}^*\mathbb{R}^{\mathbb{T}} \to {}^*\mathbb{R}^{N+1}$, $(xt)_{t\in\mathbb{T}} \mapsto \left(\frac{xNt}{T}\right)_{t\in\mathbb{R}^{N+1}}$.
Thus,

$$\bar{\mathcal{Q}}^N_{\mathbf{D}'_N} = \{Q\circ j : Q\in {}^*\mathcal{Q}^N_{\mathbf{D}'_N}\}.$$

This implies,

$$\max_{\bar{Q}\in\bar{\mathcal{Q}}^N_{\mathbf{D}'_N}} \mathbb{E}^{\bar{Q}}[{}^*\xi\circ\tilde{\cdot}] = \max_{Q\in{}^*\mathcal{Q}^N_{\mathbf{D}'_N}} \mathbb{E}^{Q}[{}^*\xi\circ\hat{\cdot}].$$

□

# Chapter 5

# Conclusion

Some of the chapters in this thesis utilize nonstandard analysis (in the sense of Robinsonian nonstandard analysis), some do not. Each chapter discusses its respective topic in detail. We will now summarize our results and conclude with potential extensions and applications.

First, in the spirit of Donsker's theorem, we proved the weak convergence of a sequence of sublinear expectations defined on a discrete state-space to a continuous-time $G$-expectation. Furthermore, we proved that for bounded continuous random variables on $\Omega$, a maximum in the representation of the $G$-expectation is attained.

Secondly, we gave an intuitive and simplified introduction to nonstandard measure theory. We constructed the extended nonstandard enlargement in terms of sequences, equivalence relations and equivalence classes with respect to binary measures. We also provided an alternative construction of the Loeb measure by establishing that an internal finitely additive measure induces a premeasure. We then extended the premeasure to a measure on the $\sigma$-algebra and we concluded by proving the uniqueness of this measure.

Thirdly, very roughly speaking, we extended the discrete time analogue of the $G$-expectation to a hyperfinite time analogue. Then, we used the characterization of convergence in nonstandard analysis to prove that the hyperfinite discrete-time analogue of the $G$-expectation is infinitely close to the standard $G$-expectation. This proof gives an alternative, combinatorially inspired construction of the $G$-expectation using nonstandard analysis and also ensures a stronger mode of convergence.

Thus, we have provided a mathematical foundation for the application of the powerful tools of nonstandard analysis to $G$-stochastic calculus.

The result of this thesis motivates several related extensions and applications that seem worth pursuing in future research:

- It would be interesting to introduce a nonstandard notion of the $G$-Itô integral in the context of the Stieltjes integral (hyperfinite sum) with respect to the hyperfinite $G$-Brownian motion, to introduce the notion of a hyperfinite $G$-martingale, and furthermore, to present an alternative proof of the martingale representation theorem under the $G$-expectation, see Soner et al. [99] and Song [102] for the formulation of the theorem in the standard $G$-framework. A simplified alternative proof of the $G$-Itô formula (see Li and Peng [69] for the conditions of the formula) for a $G$-Itô process ideally by means of Taylor expansion, as in the nonstandard proof of the classical Itô formula would be desired. Hu and Peng [56] developed the basic theory for Lévy processes under $G$-expectation also known as $G$-Lévy processes. In the spirit of Lindstrøm's [74] theory for hyperfinite Lévy processes and Lindstrøm's [75] nonlinear stochastic integrals for hyperfinite Lévy processes, we hope that the methodology in this thesis can be extended to the construction of hyperfinite $G$-Lévy processes and stochastic integrals with respect to hyperfinite $G$-Lévy processes.

- It would also be desirable to find applications of this newly developed $G$-stochastic nonstandard calculus to financial economics, especially general equilibrium theory under volatility uncertainty. Since the existence results by Radner [91] and Duffie and Shafer [41, 42] ensure that a hyperfinite incomplete financial markets economy has an equilibrium, this research enterprise appears to hold some promise, provided one can combine the nonstandard methodologies developed in Anderson and Raimondo [5] and Herzberg [53] with the equilibrium theory for volatility uncertainty obtained by Epstein and Ji [43] and Beissner [10].

  Given the importance of the theory of hyperfinite Itô integration for the equilibrium existence proof in Anderson and Raimondo [5], we hope that the notion of a hyperfinite $G$-expectation developed in this thesis may ultimately provide the mathematical foundation for both a fully-fledged nonstandard theory of $G$-stochastic integrals and also an equilibrium existence proof for continuous-time models driven by $G$-Brownian stochasticity, ideally also for the multi-agent case with trading. Another approach to this would be to extend the work by Epstein and Ji [43] to multi-agent models with trading. It would be worthwhile to reformulate and simplify Beissner's [10] equilibrium existence result using nonstandard analysis. It would also be desirable to extend the work by Herzberg [53] to establish the existence of equilibrium in a continuous-time model with a single agent (or multiple agents) in which the dynamics of the dividends follows a $G$-Lévy process.

# Bibliography


[1] Albeverio, S. and Herzberg, F. (2006). A combinatorial infinitesimal representation of Lévy processes and an application to incomplete markets. *Stochastics*, 78(5):301–325.

[2] Albeverio, S., Høegh-Krohn, R., Fenstad, J., and Lindstrøm, T. (1986). *Nonstandard methods in stochastic analysis and mathematical physics.*

[3] Anderson, R. (1976). A nonstandard representation for Brownian motion and Itô integration. *Bulletin of the American Mathematical Society*, 82:99–101.

[4] Anderson, R. (1991). Non-standard analysis with applications to economics. In *Handbook of mathematical economics. volume IV*, pages 2145–2208. Amsterdam: North-Holland.

[5] Anderson, R. and Raimondo, R. (2008). Equilibrium in continuous-time financial markets: Endogenously dynamically complete markets. *Econometrica*, 76(4):841–907.

[6] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3):203–228.

[7] Avellaneda, M., Levy, A., and Parás, A. (1995). Pricing and hedging derivative securities in markets with uncertain volatilities. *Applied Mathematical Finance*, 2:73–88.

[8] Bank, P. and Riedel, F. (2001). Existence and structure of stochastic equilibria with intertemporal substitution. *Finance and Stochastics*, 5(4):487–509.

[9] Bauer, H. (2001). *Measure and integration theory*, volume 26 of *De Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin.

[10] Beissner, P. (2013). Radner equilibria under ambiguous volatility . *Institute of Mathematical Economics Working Paper No. 493.*

[11] Bell, J. and Slomson, A. (1969). Models and ultraproducts. An introduction. Amsterdam-London: North-Holland Publishing Company 1969. X, 322 p.

[12] Berg, I. (2007). *The strength of nonstandard analysis*. Springer, Wien.

[13] Bewley, T. (1986). Knightian decision theory. Part I. *Discussion Paper, Cowles Foundation*.

[14] Bick, A. (1990). On viable diffusion price processes of the market portfolio. *The Journal of Finance*, 45(2):673–689.

[15] Billingsley, P. (1995). *Probability and measure*. Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, third edition.

[16] Billingsley, P. (1999). *Convergence of probability measures*. Wiley Series in Probability and Statistics. John Wiley & Sons Inc., New York, second edition.

[17] Bion-Nadal, J. and Kervarec, M. (2012). Risk measuring under model uncertainty. *The Annals of Applied Probability*, 22(1):213–238.

[18] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654.

[19] Bollerslev, T., Sizova, N., and Tauchen, G. (2012). Volatility in equilibrium: asymmetries and dynamic dependencies. *Review of Finance*, 16(1):31–80.

[20] Brown, D. and Robinson, A. (1975). Nonstandard exchange economies. *Econometrica*, 43(1):41–55.

[21] Carr, P. and Lee, R. (2009). Volatility derivatives. *Annual Review of Financial Economics*, 1:1–21.

[22] Chen, Z. and Epstein, L. (2002). Ambiguity, risk, and asset returns in continuous time. *Econometrica*, 70(4):1403–1443.

[23] Cheng, X. and Riedel, F. (2013). Optimal stopping under ambiguity in continuous time. *Mathematics and Financial Economics*, 7(1):29–68.

[24] Cohen, S., Ji, S., and Peng, S. (2011). Sublinear expectations and martingales in discrete time. *Preprint*.

[25] Cont, R. (2006). Model uncertainty and its impact on the pricing of derivative instruments. *Mathematical Finance*, 16(3):519–547.

[26] Cox, J., Ingersoll, J., and Ross, S. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2):363–384.

[27] Cox, J., Ross, S., and Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229–263.

[28] Cutland, N., editor (1988). *Nonstandard analysis and its applications*, volume 10 of *London Mathematical Society student texts*. Cambridge Univ. Pr., Cambridge.

[29] Cutland, N. (2000). *Loeb measures in practice: Recent advances.* Berlin: Springer.

[30] Cutland, N., Kopp, P., and Willinger, W. (1991). A nonstandard approach to option pricing. *Mathematical Finance*, 1(4):1–38.

[31] Cutland, N., Kopp, P., and Willinger, W. (1993a). A nonstandard treatment of options driven by Poisson processes. *Stochastics and Stochastics Reports*, 42(2):115–133.

[32] Cutland, N., Kopp, P., and Willinger, W. (1993b). From discrete to continuous financial models: new convergence results for option pricing. *Mathematical Finance*, 3(2):101–123.

[33] Dana, R. (1993). Existence and uniqueness of equilibria when preferences are additively separable. *Econometrica*, 61(4):953–957.

[34] Dana, R. and Riedel, F. (2013). Intertemporal equilibria with Knightian uncertainty. *Journal of Economic Theory*, 148(4):1582–1605.

[35] Debreu, G. (1954). Valuation equilibrium and Pareto optimum. *Proceedings of the National Academy of Sciences of the United States of America*, 40:588–592.

[36] Delbaen, F. (2002). Coherent risk measures on general probability spaces. In *Advances in finance and stochastics. Essays in honour of Dieter Sondermann*, pages 1–37. Berlin: Springer.

[37] Denis, L., Hu, M., and Peng, S. (2011). Function spaces and capacity related to a sublinear expectation: Application to $G$-Brownian motion paths. *Potential Analysis*, 34(2):139–161.

[38] Dolinsky, Y., Nutz, M., and Soner, M. (2012). Weak approximation of $G$-expectations. *Stochastic Processes and their Applications*, 122(2):664–675.

[39] Duffie, D. and Huang, C.-F. (1985). Implementing Arrow-Debreu equilibria by continuous trading of few long- lived securities. *Econometrica*, 53:1337–1356.

[40] Duffie, D. and Protter, P. (1992). From discrete to continuous-time finance: Weak convergence of the financial gain process. *Mathematical Finance*, 2(1):1–15.

[41] Duffie, D. and Shafer, W. (1986a). Equilibrium in incomplete markets. I: A basic model of generic existence. *Journal of Mathematical Economics*, 14:285–300.

[42] Duffie, D. and Shafer, W. (1986b). Equilibrium in incomplete markets. II: Generic existence in stochastic economies. *Journal of Mathematical Economics*, 15:199–216.

[43] Epstein, L. and Ji, S. (2013). Ambiguous volatility and asset pricing in continuous time. *The Review of Financial Studies*.

[44] Epstein, L. and Ji, S. (2014). Ambiguous volatility, possibility and utility in continuous time. *Journal of Mathematical Economics*, 50:269–282.

[45] Epstein, L. and Schneider, M. (2007). Learning under ambiguity. *The Review of Economic Studies*, 74(4):1275–1303.

[46] Epstein, L. and Wang, T. (1994). Intertemporal asset pricing under Knightian uncertainty. *Econometrica*, 62(2):283–322.

[47] Eraker, B. and Shaliastovich, I. (2008). An equilibrium guide to designing affine pricing models. *Mathematical Finance*, 18(4):519–543.

[48] Harrison, J. and Kreps, D. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408.

[49] Harrison, J. and Pliska, S. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(11):215–260.

[50] Henson, C. and Moore, L. (1974). Nonstandard hulls of the classical Banach spaces. *Duke Mathematical Journal*, 41(2):277–284.

[51] Herzberg, F. (2008). A definable nonstandard enlargement. *MLQ. Mathematical Logic Quarterly*, 54(2):167–175.

[52] Herzberg, F. (2010). Hyperfinite stochastic integration for Lévy processes with finite-variation jump part. *Bulletin des Sciences Mathématiques*, 134(4):423–445.

[53] Herzberg, F. (2013). First steps towards an equilibrium theory for Lévy financial markets. *The Annals of Finance*, 9(3):543–572.

[54] Hoover, D. and Perkins, E. (1983). Nonstandard construction of the stochastic integral and applications to stochastic differential equations. I., II. *Transactions of the American Mathematical Society*, 275:1–58.

[55] Hrbáček, K. (1979). Nonstandard set theory. *American Mathematical Monthly*, 86:659–677.

[56] Hu, M. and Peng, S. (2009). G-Lévy Processes under sublinear expectations. *Preprint.*

[57] Hurd, A. and Loeb, P. (1985). *An introduction to nonstandard real analysis*, volume 118 of *Pure and applied mathematics.* Academic Pr., Orlando.

[58] Keisler, H. (1977). Hyperfinite model theory. *Logic Colloqium.*, 76:5–110.

[59] Keisler, H. (1995). A neometric survey. In *Developments in nonstandard mathematics. Proceedings of the international colloquium on nonstandard mathematics, Portugal, 1994*, pages 233–250. Harlow: Longman; New York, NY: Wiley.

[60] Khan, M. and Sun, Y. (1997). The capital-asset-pricing model and arbitrage pricing theory: A unification. *Proceedings of the National Academy of Sciences of the United States of America*, 94:4229–4232.

[61] Klenke, A. (2014). *Probability theory: A comprehensive course.* Springer, London, 2. edition.

[62] Knight, F. (1921). *Risk, uncertainty and profit.* Houghton Mifflin, Boston, 6 edn.

[63] Kolmogorov, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer, Berlin.

[64] Kopp, P. (1997). Hyperfinite mathematical finance. In *Nonstandard analysis: theory and applications. L. Arkeryd and N. Cutland and C. Henson*, pages 279–307. Dordrecht.

[65] Kreps, D. (1981). Arbitrage and equilibrium in economies with infinitely many commodities. *Journal of Mathematical Economics*, 8:15–35.

[66] Laugwitz, D. (1961a). Anwendungen unendlich kleiner Zahlen. I: Zur Theorie der Distributionen. *Journal für die Reine und Angewandte Mathematik*, 207:53–60.

[67] Laugwitz, D. (1961b). Anwendungen unendlich kleiner Zahlen. II: Ein Zugang zur Operatorenrechnung von Mikusinski. *Journal für die Reine und Angewandte Mathematik*, 208:22–34.

[68] Laugwitz, D. (1978). Infinitesimalkalkül: Kontinuum und Zahlen-eine elementare Einführung in die Nichtstandard-Analysis. Mannheim-Wien-Zürich: Bibliographisches Institut, B.I.-Wissenschaftsverlag. 187 S.

[69] Li, X. and Peng, S. (2011). Stopping times and related Itô's calculus with $G$-Brownian motion. *Stochastic Processes and their Applications*, 121(7):1492–1508.

[70] Lindstrøm, T. (1980a). Hyperfinite stochastic integration. I: The nonstandard theory. *Mathematica Scandinavica*, 46:265–292.

[71] Lindstrøm, T. (1980b). Hyperfinite stochastic integration. II: Comparison with the standard theory. *Mathematica Scandinavica*, 46:293–314.

[72] Lindstrøm, T. (1980c). Hyperfinite stochastic integration. III: Hyperfinite representations of standard martingales. *Mathematica Scandinavica*, 46:315–331.

[73] Lindstrøm, T. (1988). An invitation to nonstandard analysis. In *Nonstandard analysis and its applications N. Cutland*, pages 1–99.

[74] Lindstrøm, T. (2004). Hyperfinite lévy processes. *Stochastics and Stochastic Reports*, 76(6):517–548.

[75] Lindstrøm, T. (2008). Nonlinear stochastic integrals for hyperfinite Lévy processes. *Logic and Analysis*, 1(2):91–129.

[76] Loeb, P. (1975). Conversion from nonstandard to standard measure spaces and applications in probability theory. *Bulletin of the American Mathematical Society*, 211:113–122.

[77] Loeb, P. and Wolff, M., editors (2000). *Nonstandard analysis for the working mathematician*, volume 510 of *Mathematics and its applications*. Kluwer Acad. Publ., Dordrecht.

[78] Łoś, J. (1955). Quelques remarques, théorèmes et problémes sur les classes définissables d álgèbres. Mathematical interpretation of formal systems. *Transactions of the American Mathematical Society*, pages 98–113.

[79] Luxemburg, W. (1969). A general theory of monads. Appl. Model Theory Algebra, Anal., Probab., Proc. Int. Sympos. Calif. Inst. Technol. 1967, 18-86 (1969).

[80] Lyons, T. (1995). Uncertain volatility and the risk free synthesis of derivatives. *Applied Mathematical Finance*, 2(3):117–133.

[81] Mas-Colell, A. and Richard, S. F. (1991). A new approach to the existence of equilibria in vector lattices. *Journal of Economic Theory*, 53(1):1–11.

[82] Müller, D. (1969). Nonstandard proofs of invariance principles in probability theory. Appl. Model Theory Algebra, Anal., Probab., Proc. Int. Sympos. Calif. Inst. Technol. 1967, 186-194.

[83] Mycielski, J. (1981). Analysis without actual infinity. *The Journal of Symbolic Logic*, 46:625–633.

[84] Ng, S. (2008). A nonstandard Lévy-Khintchine formula and Lévy processes. *Acta Mathematica Sinica. English Series*, 24(2):241–252.

[85] Nutz, M. (2013). Random $G$-expectations. *The Annals of Applied Probability*, 23(5):1755–1777.

[86] Pardoux, É. and Peng, S. (1990). Adapted solution of a backward stochastic differential equation. *Systems Control Letters*, 14(1):55–61.

[87] Peng, S. (2007). $G$-expectation, $G$-Brownian motion and related stochastic calculus of Itô type. *Stochastic Analysis and Applications*, 2:541–567.

[88] Peng, S. (2008). Multi-dimensional $G$-Brownian motion and related stochastic calculus under $G$-expectation. *Stochastic Processes and their Applications*, 118(12):2223–2253.

[89] Peng, S. (2010). Nonlinear expectations and stochastic calculus under uncertainty. *Preprint*.

[90] Perkins, E. (1981). A global intrinsic characterization of Brownian local time. *The Annals of Probability*, 9:800–817.

[91] Radner, R. (1972). Existence of equilibrium of plans, prices, and price expectations in a sequence of markets. *Econometrica*, 40:289–303.

[92] Rashid, S. (1987). Economies with many agents. An approach using nonstandard analysis. Baltimore/London.

[93] Rigotti, L. and Shannon, C. (2005). Uncertainty and risk in financial markets. *Econometrica*, 73(1):203–243.

[94] Robinson, A. (1961). Nonstandard analysis. *Nederl. Akad. Wetensch. Proc. Ser. Indag. Math.*, 23:432–440.

[95] Robinson, A. and Zakon, E. (1969). A set-theoretical characterization of enlargements. Appl. Model Theory Algebra, Anal., Probab., Proc. Int. Sympos. Calif. Inst. Technol. 1967, 109-122.

[96] Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3):341–360.

[97] Ruan, C. (2011). *The construction of G-Brownian motion and relative financial application.* Master's dissertation. School of Mathematics, Shandong University, Jinan, China.

[98] Schmieden, C. and Laugwitz, D. (1958). Eine Erweiterung der Infinitesimalrechnung. *Mathematische Zeitschrift*, 69:1–39.

[99] Soner, M., Touzi, N., and Zhang, J. (2011). Martingale representation theorem for the $G$-expectation. *Stochastic Processes and their Applications*, 121(2):265–287.

[100] Soner, M., Touzi, N., and Zhang, J. (2012). Wellposedness of second order backward SDEs. *Probability Theory and Related Fields*, 153(1-2):149–190.

[101] Soner, M., Touzi, N., and Zhang, J. (2013). Dual formulation of second order target problems. *The Annals of Applied Probability*, 23(1):308–347.

[102] Song, Y. (2011). Some properties on $G$-evaluation and its applications to $G$-martingale decomposition. *Science China Mathematics*, 54(2):287–300.

[103] Stroyan, K. and Luxemburg, W. (1976). *Introduction to the theory of infinitesimals*, volume 72 of *Pure and applied mathematics*. Academic Press, New York.

[104] Sun, Y. (2000). Economics and nonstandard analysis. In *Nonstandard analysis for the working mathematician*, pages 259–305. Dordrecht.

[105] Trojani, F. and Vanini, P. (2002). A note on robustness in Merton's model of intertemporal consumption and portfolio choice. *Journal of Economic Dynamics and Control*, 26(3):423–435.

[106] Vorbrink, J. (2014). Financial markets with volatility uncertainty. *Journal of Mathematical Economics*, 53:64–78.