APPLICATIONS OF NON STANDARD ANALYSIS TO PROBABILITY THEORY

TRISTRAM DE PIRO

Contents

1.	Introduction; Philosophy of the Method	1
2.	The Basics of Non Standard Analysis	11
3.	Loeb Integration and Measure	28
4.	Peano's Theorem	53
5.	Fourier Analysis	56
6.	The Ergodic Theorem	75
7.	Stochastic Calculus	96
References		125

1. Introduction; Philosophy of the Method

I first became interested in non-standard analysis when I was a graduate, studying for a PhD in Mathematics at M.I.T. I was introduced to the subject by my adviser Byunghan Kim, and Boris Zilber, both of whom work on the interactions between logic and geometry. Although I had not read Robinsons' foundational book on the subject at the time, see [36], I was looking at Zariski geometries, which employ infinitesimals indirectly to define a tangency relation between curves defined in logical structures. This is a crucial ingredient in a Trichotomy Theorem due to Zilber and Hrushovski, see [20], which states roughly that any such curve has a very simple geometry attached to it, the trivial case, or interprets a group with no extra structure, the so called one-based case, or interprets a field. This theorem was later used by Hrushovski to prove several powerful results in algebraic geometry and number theory, namely the Mordell Lang conjecture, see [18], and the Manin Mumford conjecture, see [19]. Since then, the method of infinitesimal neighborhoods, has been clarified greatly by Zilber in [45], and various applications of the trichotomy have been found in number theory by Thomas Scanlon, see his webpage http://math.berkeley.edu/ scanlon.

My PhD thesis was mainly concerned with one-based geometries in so called simple theories, and generalisations of the tangency method I mentioned above. However, due to the high degree of abstraction involved, I later found it more useful to work in more concrete algebraic structures and tried to reprove a classical result in the theory of algebraic curves, "Bezout's Theorem", using nonstandard methods. This I managed to do, while I was a postdoc at Edinburgh University. The results of this research can be found in [32]. Essentially Bezout's theorem says that the number of intersection points of two algebraic curves, including multiplicity, is equal to the product of the degrees of the curves. However, the classical proof relies on a purely algebraic definition of multiplicity, and in many ways, is fairly non-intuitive. The definition involving infinitesimal neighborhoods, defines multiplicity by counting the number of points of intersection, close to the point in question, after a small variation of the curves. By proving the equivalence of this definition with the algebraic one, the proof of Bezout's theorem becomes almost trivial. It is sufficient to observe that two *generic* curves in the family intersect in de points, where d and e are the degrees of the curves. This is clear as the intersections are transverse simple points. I later discovered that this was in fact the *original* method of proving the theorem, due to Poncelet, however, due to its reliance on the intuitive non-rigorous notion of "Conservation of Number", it was replaced by an algebraic proof by Zariski. The desire to find a rigorous justification of "Conservation of Number" arguments is, in fact, part of Hilbert's 15th problem which is still to be resolved. The method of infinitesimal neighborhoods was the cornerstone of the "Italian School of Algebraic geometry", which included Francesco Severi, Castelnuovo and Enriques. There is no doubt that, although many of their arguments are not rigorous by modern standards, their intuitions on curves and surfaces laid the foundations of modern algebraic geometry. I went on to study some of Severi's work, while a postdoc at Camerino University in Italy, and produced a number of papers which justify his intuitive arguments, using the infinitesimal method, mainly in [38]. These papers can be found on my website http://www.magneticstrix.net.

Of course, as even most non-mathematicians know, infinitesimal quantities were discovered, possibly independently, by Newton and Liebniz, and resulted in the development of calculus. Unfortunately, Liebniz published his results in 1684, while Isaac Newton didn't give a full account of his version of the calculus until 1704, in [26]. Newton's circle of friends claimed that he had obtained his results earlier, around 1667, and before Liebniz, which led to an acrimonious dispute, eventually accusing Liebniz of plagiarism, based on his having obtained a copy of Newton's manuscript [27] in 1675. As I argue in [33], the paper

[27] does not give a clear exposition of the "method of fluxions", and lacks a real proof of the Fundamental Theorem of Calculus, although it does give an intuitive definition of integration $(*)$, so the accusation against Liebniz is probably unjustified. However, consideration of an earlier manuscript [28], which does present a proof of The Fundamental Theorem of Calculus, and [29], which shows that Newton was familiar with differentiation, or the method of tangency, support the view that Newton had formulated his own version of the calculus by 1669. Newton's early unpublished papers are still interesting to read, and I tried to formulate a rigorous version of his arguments, especially his early proof of the Fundamental Theorem of Calculus, in [33]. Newton's definition of integration, $(*)$, using infinitesimals, can be given a rigorous formulation, which we will consider in Section 2.

The methods of both Liebniz and Newton were famously criticized by George Berkeley in [5]. There is a logical paradox in Newton's calculus, which require that an infinitesimal quantity ϵ should be both zero and non-zero at different stages of his arguments. The rigorous formulation of the calculus in the 19th century was able to resolve these logical problems by replacing the notion of an infinitesimal quantity with that of a limit, however, in doing so, it perhaps changed the intuitive character of the calculus. It was not until Abraham Robinson's pioneering work [36], that infinitesimals were given a solid foundation. We will consider Robinson's argument more closely in Section 2. I think it is fair to say, that, even now, Robinson's work is unpopular due to its reliance on non first order logic. This seems to be rather unfortunate, as the existence of infinitesimals can be easily shown in the first order context, and the introduction of non first order logic, which is necessary for more advanced arguments, seems to be a small price to pay, in order to apply what seems to be an extremely powerful method, which I will now discuss.

The method that I wish to consider and which is applied in this dissertation, goes beyond the geometric considerations that I have outlined above, and introduces infinitesimals not just as a way of making standard arguments in mathematics more intuitive, but by finding a new strategy of proof. The method essentially consists in trying to reduce problems concerning the infinite to the finite case, which is probably easier to solve. The idea is roughly as follows;

First Step; Formulate a finite version of the theorem you want to prove, and show it in this case. Sometimes the reformulation of the problem to the finite case leads to a trivial solution, in which case the method probably won't work. If this fails, then try to find some part of the original problem, which leads to an interesting solution in the finite case, and then follow the remaining steps to see if this helps to solve the problem.

Second Step; Transfer the results of the finite case to the hyperfinite, this can be done using the logical methods developed by Robinson. We will discuss them in Section 2.

Third Step; Use the theory of Loeb integration and measure to specialise the results from the second step, to a standard space, close to the original problem. We will discuss this theory in Section 3.

Fourth Step; Usually using standard methods, find a way of relating your resulting standard space to the original space in question, and thus prove the theorem.

This thesis is organised as follows;

In Section 2, we consider the logical foundations of Robinson's method, which allows you to carry out the second step, this work is mainly based on [36] and the logical notes in [12]. In Section 3, we consider the mechanics of Loeb integration and measure, based mainly on the paper by [24] and the notes in [12]. The extensions of Loeb's method are mainly due to myself, though references can be found in [12]. This technique allows you to carry out the third step. Steps 1 and 4 depend on the problem in question, so we discuss the theorems required to carry out this part in the course of discussing each application. We illustrate the method in four important cases in probability theory and analysis;

Case 1(Section 4) Peano's Theorem.

This is very short section, and is based on the proof in [12]. The method originates in [22]. The first step just consists in the observation that one can solve difference equations over finite sets, recursively. Then, using the second step, one can find solutions to hyperfinite difference equations, on a hyperfinite interval. The third step is technically the most difficult and relies on a lifting result in [2], for functions taking values in a Hausdorff space, with a countable base of open sets. It also requires the general theory of Loeb integration, which we develop in Section 3. The final step is based on the observation, that the standard part mapping of the hyperfinite interval onto the standard interval, is measurable and measure preserving, we prove a more general result in

Section 5, see also Section 7. In particular, it follows, by change of variables, that we can relate the Loeb integral from the third step, to a standard integral.

Case 2(Section 5) Fourier Analysis.

In this section, we prove the Inversion Theorem for Fourier transforms on the real line. I am not aware of any published result of this theorem using nonstandard methods, so it perhaps constitutes the most original part of the thesis. The first step consists in proving an analogue of the theorem for finite groups. More precisely, we prove a character formula, based on results in [25]. The second step involves a purely logical argument, which transfers this result to a relation of hyperfinite sums. This can be converted to a result involving internal integrals. We consider both these notions in Section 2. The third step specialises the internal integrals to Loeb integrals on an unbounded hyperfinite real interval, using the standard part mapping. This step is technically the most difficult as it involves the verification of certain conditions on the integrals of the tails of the internal functions concerned, $(*)$, these conditions are discussed in Section 3. In order to obtain these conditions, we assume that the original function f belongs to the Schwarz class, though these assumptions can be relaxed. We can then easily obtain the required condition for the transfer of f, done in Section 5. In order to satisfy the condition for the hyperfinite analogue of the Fourier transform $\hat{f}$ of f $(**)$, we study the decay rate of discrete Fourier transforms of f. This involves introducing a basic calculus around discrete Fourier transforms, done in Section 5. More precisely, we prove a fundamental theorem of calculus for discrete derivatives, a simple result relating the discrete fourier transforms of shifts of a discrete function, a product formula for derivatives, and an analogue of the classical result relating the Fourier transform of a function and its derivative. We then show that the discrete Fourier transforms of f and its derivatives are uniformly bounded, in Section 5, and proceed to the integral calculation which verifies the condition $(**)$. (I originally tried to obtain the condition $(**)$ using the theory of lifts in Loeb integration, but could find no way of relating the resulting functions using the hyperfinite analogue of the Fourier transform.) Again, there are some technical problems in transferring this estimate $(**)$, which we then consider. The application of the Loeb technique essentially involves taking standard parts of both sides of the equation we obtained by transfer from the finite case. The verification of the condition $(**)$

ensures that the relation is preserved for the specialised functions on the hyperfinite reals. This calculation is carried out at the end of Section 5. Finally, we carry out the fourth step, which involves showing that the standard part mapping is a measurable, measure preserving map from the unbounded hyperfinite real interval to the reals with the addition of $\{+\infty, -\infty\}$, $(***)$. The result then follows from a simple change of variables argument, for ease of exposition, we consider this argument at the beginning of Section 5. An almost identical step to $(***)$ is used in the proof of Peano's theorem, and the stochastic calculus arguments. However, in these cases we only need the result for bounded hyperfinite real intervals, which follows easily from the more general case considered in Section 5. In this section, I relied heavily on the texts [23], [40] and [3].

Case 3(Section 6) The Ergodic Theorem.

In this section, we prove the ergodic theorem for probability spaces. The proof is based on the paper [21]. Again we follow the same basic method. Unfortunately, the content of the the ergodic theorem for finite probability spaces is trivial, so the first step, doesn't work directly, as in Section 5. Instead what can be shown for finite probability spaces, which are ordered by a measure preserving transformation, is how to relate the averages of two functions F and G for which there exists a simple relationship on subintervals in the ordering. The second step transfers this idea using the method of hyperfinite sums, as before, to hyperfinite intervals. In the third step, Loeb integration theory is used to approximate the function f appearing in the statement of the ergodic theorem, and its limsup $\overline{f}$, by the functions F and G. This approximation lemma is just stated without proof in the paper, but I have supplied this at the end of Section 3. A simple local relationship holds between f and $\bar{f}$, on a hyperfinite interval, and this property is inherited by the approximations F and G. Using the second step, we obtain, as before, a relationship between the internal integrals, defined in terms of hyperfinite sums, $(*)$. The main technical difficulty here, is to bound the remainder term coming from the partitioning of a hyperfinite interval into subintervals by a standard finite number. Kamae is able to find a simple solution to this problem using a simple compactness or overflow argument, I explain these notions in Section 2. The third step, specialises the relationship $(*)$ to obtain an inequality between the integral of f and $\bar{f}$, on the hyperfinite interval. Using the fact that the remainder term obtained in the second step is finite, and thus has Loeb measure zero, the inequality holds over the whole space. Of course, we encountered a similar problem in Section 5, though, in this case, it seems to be resolved by analytic rather than logical arguments. This essentially proves the ergodic theorem for hyperfinite intervals. The rest of Kamae's paper is concerned with the fourth step, and addresses the question of how to obtain a proof of the ergodic theorem for arbitrary probability spaces from a proof for hyperfinite ones. This construction uses entirely standard methods. The idea is to show, first, that the ergodic theorem for arbitrary probability spaces follows from the theorem for $\mathcal{R}^{\mathcal{N}}$, the reals indexed by the natural numbers, considered with the Borel field and a shift invariant measure, $(*)$. Second, one shows that the space obtained in $(*)$ is isomorphic, as a dynamical system to $[0,1]^{\mathcal{N}}$, again considered with the Borel field, and a shift invariant probability measure, $(**)$. Third, one proves that the space in $(**)$ is a factor of a hyperfinite probability interval, $(***)$. This proves the theorem for the space considered in $(*)$, as once you have shown $(**)$ and $(***)$, you can deduce the theorem for $(*)$, from the proof for hyperfinite spaces, which we have shown. The arguments $(*)$ and $(**)$ are both straightforward. The difficult part is to show $(***)$. This follows by proving the existence of "typical" elements, for which one *does not* need the ergodic theorem, to construct the factor map from the hyperfinite interval to the space in $(**)$. The existence of "typical" elements for a shift invariant measure, was apparently originally proved by de Ville, although I haven't seen his paper. In Kamae's paper, he first claims that this follows from the convergence of periodic measures to a shift invariant measure, though he doesn't give a proof of this claim. I have supplied this in the section. The proof of the convergence of periodic measures requires the formulation of a criteria in terms of box measure, using the Riesz Representation Theorem and some simple functional analysis, again I have supplied this. Kamae shows this criteria is satisfied by periodic measures using an ingenious counting argument, though I have clarified many parts of his proof, including the graph theory part of his argument, see footnote 22. As well as the paper, I relied mainly on [37].

Case 4(Section 7) Stochastic Calculus.

This section is based on the paper [1]. I also relied on [39] and [44]. We consider three main problems;

(i). The proof of the existence of Brownian motion.
(ii). A definition of stochastic integration, with respect to Brownian

motion, and its property of being a martingale up to stochastic equivalence.
(iii). A generalisation of Ito's Lemma.

All three problems are solved using the basic method outlined above. We consider each problem in turn.

(i). In the first step, we use the *standard* Central Limit Theorem for hyperfinite, but see $(*)$ below, Loeb probability spaces. These are just sequences of coin tosses indexed by a hyperfinite integer, with the specialised counting measure,and inherit the basic properties of spaces involving a finite number of tosses. Even though such Loeb spaces are not definable within the reals, one can interpret the theorem as a standard relationship using the relevant distribution function G and the Gaussian Ψ. In the second step, we transfer this result to obtain the distribution of hyperfinite sums of what are referred to as $*$-independent random variables. The transfer argument relies on a general lemma that the specialised random variables are independent with respect to Loeb measure, $(**)$. Under the assumption that each *-independent random variable has a distribution function F which is the transfer of a standard function, one can show that the specialised variables are also identically distributed, with distribution G, and that F is the transfer of G. The usual logical arguments then tranfer the Central Limit Theorem to a statement about the distribution of hyperfinite sums of *-independent random variables in terms of the transfer $^*\Psi$ of the Gaussian. $(*)$, One might be able to transfer the statement of the Central Limit Theorem for finite measure spaces directly, as by the Berry Essen theorem, the rate of convergence depends only on the moments of the random variables involved rather than the size of the measure space. This would simplify Anderson's argument considerably, as well as avoiding the use of specialisations, which usually appear in the third step. I hope to consider this problem in more detail at a later stage. The result of this transfer argument becomes a remarkable property of the distribution of random walks, with an infinite number of steps, on hyperfinite probabilty spaces with counting measure, as such random walks are a special case of *-independent random variables, by transfer of the result for finite measure spaces. In the third step, Anderson specialises a hyperfinite random walk to Brownian motion on the Loeb space, using the standard part map. This does not involve any Loeb integration theory, just a simple relationship between hyperfinite counting measure and its specialisation. The verification of the Gaussian property for the increments of Brownian motion, follows from that of infinite increments in the random walk, and the independence between increments follows easily from $(**)$. The fourth step is currently unresolved, it is still an open question as to whether you can construct Brownian motion on arbitrary probability spaces this way. Perhaps one can adapt the steps in Kamae's argument.

(ii). The first part of the problem is resolved in Anderson's paper, I managed to solve the second part myself. I will consider mainly the steps in Anderson's argument here. As usual, we need to transfer a result from finite measure spaces. This is the existence of the Stiejtes integral, and transfers to a $*$-Stiejtes integral on a hyperfinite space with respect to the increments of a hyperfinite random walk. In the third step, which is, as usual, the technically most difficult, Anderson proves the existence of liftings of standard integrable functions on $\mathcal{L}(A) \times [0,1]$, to $A \times {}^*[0,1]$, where $\mathcal{L}(A)$ is the Loeb space associated to the hyperfinite space A of coin tosses. Such liftings satisfy the Loeb integration criteria, and allow the specialisation arguments to go through. In order to prove the existence of such liftings, we first lift the function in question on the interval $[0,1]$ to the hyperfinite interval ${}^*[0,1]$. This relies on a theorem which we prove in more generality in Section 5. Using general Loeb integration theory, see Section 3, and some facts on product measure, one can then lift to the product space. An argument using expectations ensures that the lifted functions also satisfy good measurabilty properties on the hyperfinite space A, with respect to the counting measure ν. (I initially tried to adapt this argument in Section 5, but was unable to relate the lift of f and its Fourier transform $\hat{f}$.) Anderson then applies the definition of a *-Stiejtes integral to the lifted function. The analogy with the standard definition of the stochastic integral is quite revealing. It is necessary to check that this is a good definition, by showing that two different liftings produce stochastically equivalent processes, as in the standard case. Finally, one checks that the definition coincides with the standard one, by verifying the result for functions which are locally constant on $[0,1]$. This involves a lifting argument which is not explained in the paper, but I clarify it here. The fourth step of course depends on the resolution of the problem in (i). For the martingale argument, I essentially use the fact that random walks are martingales in the first step, and obtain the corresponding result for *-Stiejtes integrals on A, in the second step, by a simple calculation. I specialise this property in the third step to the stochastic integrals, but there are some technical problems which I resolved using

Ito's isometry.

(iii). This generalisation states that a function h in m variables of a sum of stochastic and non-stochastic integrals, $(*)$, with respect to n independent Brownian motions, is again such a sum, and gives an explicit formula in terms of the original integrands, and derivatives of h. In the paper, the case when $m = 1$ and $n = 2$ is shown. For variety, we show the case $m = 2$ and $n = 1$. Any other case can be reduced easily to these. In the first two steps, we transfer the standard result of Taylor's theorem for functions on the real line in two variables, and, as before, use the independence properties of *-finite random walks, and the *-finite sum representation of internal and *-Stiejtes integrals. Of course, as with all standard theorems, one can try to prove them using the nonstandard approach, so the introduction of Taylor's theorem doesn't contradict the basic idea of the method. We use these properties to calculate $\theta(t, \omega) - \theta(0, \omega)$, where θ is the transfer of $h(G(t, \omega))$ and $G(t, \omega)$ is the lift of the combination in $(*)$. As in the standard proof, we write this as an alternating sum, and use the transfer of Taylor's theorem, together with the *-independence of increments in the random walk, and the fact that the square of the increments is the horizontal step size of the random walk, an infinitesimal. This last simple observation is perhaps the main advantage over the standard proof. In the third step, we specialise the results of this calculation. The main difficulty is in relating the specialised terms to the appropriate integrals on the Loeb space. In the case of Ito stochastic integrals, this follows easily from the construction in (ii). In the non-stochastic case, one relies on a technical lemma, involving Loeb integrals, and Fubini's theorem.

It has been a great pleasure to work on this thesis, and discover the elegance of this method. In principle, one could hope to apply it in many different areas of mathematics, either to reprove standard results, or discover new ones. The area of p-adic integration which is connected to problems in algebraic geometry, seems a particularly promising area of study, as it is already phrased in terms of measures. One might even hope to find a connection between calculations over finite fields, such as the Weil conjectures, with calculations over number and function fields this way. However, I can only speculate on how this might be done. I would like to thank the people who have helped me in writing this thesis, my advisers Dominic McCarthy and Mark Holland, as well as Nigel Cutland, Teturo Kamae, Robert Anderson, Richard Kaye and

Mitchell Berger. Most of all, I would like to thank my mother and my sister, for their constant encouragement and support.

2. The Basics of Non Standard Analysis

In this section, we introduce the basic principles of non-standard analysis. The main idea is to construct an extension of the reals $\mathcal{R}$, which contains infinitesimal, or infinitely small, non-zero elements. This extension is usually dentoted by ${}^*\mathcal{R}$, and is referred to as the hyperreals. We make the following definition, see also [12].

Definition 2.1. *Let $x, y \in {}^*\mathcal{R}$, we say that;*

(i). x is infinitesimal, if, $|x| < \epsilon$, for all $\epsilon > 0$, with $\epsilon \in \mathcal{R}$.

(ii). x is finite if there exists $r \in \mathcal{R}$, with $|x| < r$.

(iii). x is infinite if $|x| > r$ for all $r \in \mathcal{R}$.

(iv). x is infinitely close to y, denoted $x \simeq y$, if $x - y$ is infinitesimal.

If $x \in \mathcal{R}$, we define the monad $\mu(x) = \{z \in {}^*\mathcal{R} : x \simeq z\}$.

Remarks 2.2. *The modulus function is just the same as for the reals, namely $|x| = x$, if $x \geq 0$, and $|x| = -x$ if $x < 0$. The terminology of a monad is due to Robinson (he was obviously a follower of Liebniz). In the theory of Zariski structures, which we briefly referred to in the introduction, a monad is intuitively the same concept as an infinitesimal neighborhood. However, the definitions are not the same, and it requires a little effort to show that the two definitions are equivalent in, for example, the case of the complex numbers $\mathcal{C}$ and the hyper complex numbers ${}^*\mathcal{C}$, with the complex topology. I showed this in the paper* [34].

There are two methods available to construct ${}^*\mathcal{R}$, one can use a compactness argument or an ultraproduct construction. I will show the first method here, the ultraproduct construction can be found in [12]. I assume some familiarity with first order logic, but will explain some of the notions, in the remark at the end of the proof.

Lemma 2.3. *Let $\mathcal{L} = < 0, 1, <, +, . >$ be the language of ordered fields. Let $\mathcal{R}$ be the reals, considered as a first order structure in this language. Let R be a set of constants, denoting elements of $\mathcal{R}$, and let c be a new* *constant symbol. Then the following theory is consistent;*

$$T = Th(\mathcal{R}, R) \cup \{0 < c < r : r \in R, r > 0\}$$

Let ${}^*\mathcal{R}$ be a model of this theory T, then ${}^*\mathcal{R}$ is an ordered field, is an elementary extension of $\mathcal{R}$, and contains non-zero infinitesimal elements.

Proof. By the compactness theorem, see [16], to show the first part, it is suffient to verify that any finite subset of the sentences in T is consistent. Without loss of generality, such a subset can be written as the single sentence;

$$\sigma = \phi(r_1, \ldots, r_n, r_{n+1}, \ldots r_m) \wedge 0 < c < r_1 \wedge \ldots \wedge 0 < c < r_n$$

where $\phi(r_1, \ldots, r_m)$ denotes a true sentence in the structure $\mathcal{R}$, involving the constant symbols $\{r_1, \ldots, r_m\}$. As the constant symbols $\{r_1, \ldots, r_n\}$ all denote elements of $\mathcal{R}$ which are greater than zero, we can clearly interpret the new constant symbol c in $\mathcal{R}$ to satisfy the sentence $0 < c < r_1 \wedge \ldots \wedge 0 < c < r_n$, so $\mathcal{R}_c$ satisfies σ. For the second part, let ${}^*\mathcal{R}$ be a model of the theory T, then ${}^*\mathcal{R}$ is an ordered field, as the axioms of an ordered field are contained in $Th(\mathcal{R}, R) \subset T$. The elementary extension property follows easily from the definition of $Th(\mathcal{R}, R)$, see the remark below. Finally, the interpretation $c^{{}^*\mathcal{R}}$ is infinitesimal, by construction.

□

Remarks 2.4. *Given two first order structures $\mathcal{M}$ and $\mathcal{N}$, we say that $\mathcal{N}$ is an elementary extension of $\mathcal{M}$, written as $\mathcal{M} \prec \mathcal{N}$, if $M \subseteq N$ and for every formula $\phi(\bar{x})$ in the language, involving elements $\bar{m}$ from $\mathcal{M}$;*

$$\mathcal{M} \models \phi(\bar{m}) \text{ iff } \mathcal{N} \models \phi(\bar{m}).$$

This is what Robinson refers to as a transfer principle, it allows you to transfer any sentence true in the given language, from $\mathcal{R}$ to ${}^*\mathcal{R}$.

Given a first order structure $\mathcal{M}$, in a language $\mathcal{L}$, and a set of constants M, denoting the elements of $\mathcal{M}$, by the theory $Th(\mathcal{M}, M)$, we mean the set of sentences of the form $\phi(\bar{m})$, (where $\phi(\bar{x})$ is a formula in the language $\mathcal{L}$, and $\bar{m}$ is a tuple of elements from M, substituted for the free variables $\bar{x}$), which are true in $\mathcal{M}$. It is an easy exercise *to show that any model* $\mathcal{N}$ *of* $Th(\mathcal{M}, M)$ *is an elementary extension of* $\mathcal{M}$.

Once you have constructed ${}^*\mathcal{R}$ this way, it is easy to see that the structure contains lots of infinitesimal elements. For example any positive integer power of an infinitesimal c is also infinitesimal. Referring back to Definition 2.1, it clearly contains finite elements, any element of $\mathcal{R} \subset^* \mathcal{R}$ is finite. It also contains infinite elements, just take the reciprocal or inverse of any non-zero infinitesimal. Every element x from $\mathcal{R}$ has a non-trivial monad, take an infinitesimal c and consider $x + c$. It is worthwhile being familiar with the basic arithmetic of infinitesimals. Of great importance is the notion of a standard part mapping or specialisation. This relies on the following simple lemma, whose proof can be found in [12] or [35], see also the brief discussion of topology at the end of this section.

Lemma 2.5. *Let* $x \in {}^*\mathcal{R}$ *be finite, then there exists a unique* $r \in \mathcal{R}$ *such that* $x \in \mu(r)$.

We make the following definition, see also Section 5, Theorem 5.7;

Definition 2.6. *The standard part mapping* $\circ : {}^*\mathcal{R} \rightarrow \mathcal{R} \cup \{+\infty, -\infty\}$

is given by setting ${}^\circ x = r$, *if* x *is finite, where* r *is the element provided by the previous lemma,* ${}^\circ x = +\infty$, *if* x *is infinite and positive, and* ${}^\circ x = -\infty$, *if* x *is infinite and negative.*

Unfortunately, this simple construction of ${}^*\mathcal{R}$ is insufficient for our purposes. We introduce the following notation;

Definition 2.7. *We define a set of types* T *according to the rule;*

*(i).*0 *is a type.*

(ii). If n *is a positive integer and* $\tau_1, \ldots, \tau_n$ *are types, then* $(\tau_1, \ldots, \tau_n)$ *is a type.*

For a set A, *we define* $V_0(A) = A$, *and, inductively define* $V_\tau(A) = P(V_{\tau_1}(A) \times \ldots \times V_{\tau_n}(A))$, *where* $\tau = (\tau_1, \ldots, \tau_n)$, *and* P *denotes the powerset operation. We let* $V(A) = \bigcup_{\tau \in T} V_\tau(A)$.

We need to work in a bigger universe which allows us to transfer statements not just about elements of $\mathcal{R}$, but about elements of $V(\mathcal{R})$.

Moreover, we need to be able to quantify over such elements. This takes us beyond the realm of first order logic. In [12], a construction called Mostowski collapse is used to produce a nonstandard universe for which such a strong transfer principle holds, however, I am not very familiar with this method. In [36], Robinson introduces what are referred to as higher order structures $\mathcal{M}$, based on a set A. Roughly speaking, the individuals of such a structure are elements of the universe $V(A)$, but we *do not* require that every individual is named. The quantifiers range over this restricted set of elements which we call *internal*. We require that the elements of A themselves are internal. We work in a fixed typed language $\mathcal{L}$ containing a relation symbol ϕ_τ for each type $\tau \neq 0$, where, if $\tau = (\tau_1, \ldots, \tau_n)$, then ϕ_τ should be $n+1$-ary, and interpreted as the membership relation between elements in $V(A)$ of compatible type. It is extremely important to realise that the internal elements do not constitute the whole of the universe $V(A)$. It is easy to show that the existence of a structure $< {}^*\mathcal{R} >$, based on a set ${}^*\mathcal{R}$ containing the reals $\mathcal{R}$, for which;

(i). There are non-zero infinitesimals.
(ii). A transfer principle holds.
(iii).The internal elements are all of $V({}^*\mathcal{R})$

leads to a contradiction;

Take an infinitesimal $\epsilon > 0$, then we would have $\frac{1}{\epsilon} > r$, for any $r \in \mathcal{R}$. Consider the set $W = {}^*\mathcal{R}_{>0} \setminus \mathcal{R}$, which contains $\frac{1}{\epsilon}$. If $\omega \in W$, then $\omega - 1 \in W$, as if $\omega - 1 \in \mathcal{R}$, then $\omega \in \mathcal{R}$. Therefore;

$$<^* \mathcal{R} > \models \sigma$$

where σ is the sentence $\exists x \forall y(\phi_{(0)}(x, y) \rightarrow y > 0 \wedge \phi_{(0)}(x, y - 1))$.

By the transfer principle;

$$< \mathcal{R} > \models \sigma$$

Hence, there exists a set $W \subset \mathcal{R}$, bounded below by 0, with the property that if $y \in W$, then $y - 1 \in W$. Clearly, no such set can exist.

The contradiction arises because W is not an internal set, we call such sets external. With the restriction on quantifiers in place, Robinson shows, essentially using the compactness theorem, that there exists a structure $< {}^*\mathcal{R} >$, satisfying a transfer principle over the reals, containing infinitesimal elements,([1]). The transfer principle takes the following form;

Theorem 2.8. *Transfer Principle*

Let $< {}^*\mathcal{R} >$ be as above, see also the following footnote. Then, if σ is a sentence in $\mathcal{L}_R$, the language obtained from $\mathcal{L}$ by adding constants for every *element of $V(\mathcal{R})$. Then;*

$$< {}^*\mathcal{R} > \models \sigma \textit{ iff } < \mathcal{R} > \models \sigma$$

In particular, for any for any sentence of the form $\phi(\bar{c}, \bar{f}, \bar{R})$, where $\{\bar{c}, \bar{f}, \bar{R}\}$ are standard elements, functions and relations in $V(\mathcal{R})$, we have that;

$$< \mathcal{R} > \models \phi(\bar{c}, \bar{f}, \bar{R}) \textit{ iff } < {}^*\mathcal{R} > \models \phi({}^*\bar{c}, {}^*\bar{f}, {}^*\bar{R})$$

Proof. The idea is to include the theory $Th(< \mathcal{R} >, R)$ when applying compactness, see Lemma 2.3.

□

A basic problem in non-standard analysis is to determine which elements of $V({}^*\mathcal{R})$ are internal. It follows from the definition of a higher order structure that every element of an internal set is internal, and every element of ${}^*\mathcal{R}$ is internal. As all the elements of $V(\mathcal{R})$ are named, the *transfer* or interpration ${}^*R$ in $< {}^*\mathcal{R} >$ of an element $R \in V(\mathcal{R})$ is internal. Note that this interpretation does *not* coincide with the

[1] Robinson's construction holds for any structure $< \mathcal{M} >$. Technically he produces a structure $< \mathcal{N} > \supset < \mathcal{M} >$, which satisfies all concurrent relations on $< \mathcal{M} >$, as well as all the sentences true in $< \mathcal{M} >$, within the language $\mathcal{L}_M$, obtained by adding constants for *every* element of $V(M)$. He calls such such an extension an enlargement, $(*)$. The concurrent relations are just relations which are finitely satisfiable in $< \mathcal{M} >$. We can obtain such structures easily using compactness. The property $(*)$ in the case of the structure $< {}^*\mathcal{R} >$ is stronger than the existence of infinitesimals and a transfer principle, as the relation $0 < x < y$ is concurrent, see the proof of Lemma 2.3. However, this extra strength is rarely used in applications. Nevertheless, we will assume that $< {}^*\mathcal{R} >$ is an enlargement of $< \mathcal{R} >$ in this sense. The reader should look at [36] for more details.

interpretation of R in $< \mathcal{R} >$, unless R is a finite set! This follows from the definition of an enlargement, see [36] Theorem 2.11.2. The following lemma is immediate;

Lemma 2.9. *Suppose that $\phi(\bar{x})$ is a formula in the language $\mathcal{L}_{int}$, the language obtained from $\mathcal{L}$ by adding constants for all internal elements of $V({}^*\mathcal{R})$. Then ϕ can be written as $\psi(\bar{x},\bar{a})$, where $\psi(\bar{x},\bar{y})$ is a formula in the language $\mathcal{L}_R$, and $\bar{a}$ are constants corresponding to internal elements of $V({}^*\mathcal{R})$, not in $V(\mathcal{R})$.*

Proof. Just substitute all the internal elements, not appearing in R, by free variables.

□

The following lemma is useful;

Lemma 2.10. *Internal Definition Principle*

Suppose that $D(\bar{x})$ is internal, of type τ, and $\psi(\bar{x},\bar{a})$ is as above. Then $D(\bar{x})\cap\psi(\bar{x},\bar{a})$ is internal. In particular, the solution set of $\psi(\bar{x},\bar{a})$ is internal.

Proof. Let σ be the sentence $\forall\bar{w}\forall\bar{x}\forall z\exists t((\phi_\tau(z;\bar{x})\wedge\psi(\bar{x},\bar{w}))\leftrightarrow\phi_\tau(t;\bar{x})$.

Then;

$< \mathcal{R} >\models \sigma$

and by the transfer principle;

$< {}^*\mathcal{R} >\models \sigma$

In particular, as $D(\bar{x})$ is internal, named by d, and $\bar{a}$ is a tuple of internal elemants, then, by the restriction on quantifiers, we can find an internal t_0, such that;

$< {}^*\mathcal{R} >\models \sigma'$

where σ' is the sentence $\forall\bar{x}((\phi_\tau(d;\bar{x})\wedge\psi(\bar{x},\bar{a}))\leftrightarrow\phi_\tau(t_0;\bar{x})$.

As every element of D and t_0 is internal, then $D(\bar{x})\cap\psi(\bar{x},\bar{a})$ consists exactly of the elements satisfying the internal relation t_0. The final part of the lemma follows by taking $D(\bar{x})$ to be $({}^*\mathcal{R})^n$, where n is the length of the tuple $\bar{x}$.

□

From now on, we will simplify notation slightly and use the symbol ${}^*\mathcal{R}$ to denote both the logical structure $< {}^*\mathcal{R} >$ and its domain. We will use mathematical notation, rather than logical notation, when writing sentences in the language $\mathcal{L}_{int}$, hoping this will not cause any confusion. We denote by ${}^*\mathcal{N} \subset {}^*\mathcal{R}$ the transfer of the natural numbers $\mathcal{N}$. We note the following;

Lemma 2.11. *The following sets are external;*

(i). ${}^*\mathcal{N} \setminus \mathcal{N}$. *(ii).* ${}^*\mathcal{R} \setminus \mathcal{R}$. *(iii).* $\mathcal{N}$. *(iv).* $\mathcal{R}$.

Proof. We just prove (i), (iii) then follows immediately, as the complement of an internal set is internal, using the transfer principle. (iv) is then immediate, as, similarly, the intersection of two internal sets is internal, and $\mathcal{R} \cap {}^*\mathcal{N} = \mathcal{N}$. Then (ii) follows from the same reasons as (iii). To show (i), observe that ${}^*\mathcal{N} \setminus \mathcal{N}$ is non-empty and has no least element. It is easy to formulate a sentence in the language of $\mathcal{R}$ saying that every nonempty subset of $\mathcal{N}$ has a least element. By the transfer principle, this statement holds for all internal subsets of ${}^*\mathcal{N}$. Hence, ${}^*\mathcal{N} \setminus \mathcal{N}$ cannot be internal.

□

An important consequence of this lemma are the following principles;

Lemma 2.12. *(i).(Overflow). Let* $A \subset {}^*\mathcal{R}$ *be an internal set;*

If there exists $m \in \mathcal{N}$, *such that for all* $n \in \mathcal{N}$, *with* $n \geq m$, $A(n)$ *holds, then there exists an infinite natural number* $\omega \in {}^*\mathcal{N}$ *such that* $A(\omega')$ *holds for all* $m \leq \omega' \leq \omega$, *([2]).*

---

[2]In [12], the overflow principle is stated in the following form; if A contains arbitrarily large finite numbers then it contains an infinite number. However, this version is generally not very useful. The proof follows immediately from the lemma, let $C = \{x \in {}^*\mathcal{N} : \exists y(A(y) \wedge y > x)\}$, then C is internal and contains $\mathcal{N}$, hence, it contains an infinite ω. There is also an in underflow principle; if A contains arbitrarily small positive infinite numbers then it contains a positive finite number. The proof is an easy exercise.

(ii). (Internal Induction). Let $A \subset {}^*\mathcal{N}$ *be an internal set;*

If for every $\omega \in {}^*\mathcal{N}$, $A(\omega)$ *implies* $A(\omega+1)$, *and* $A(0)$ *holds, then* $A = {}^*\mathcal{N}$.

Proof. (i). Let;

$$B = \{x \in {}^*\mathcal{N} : (0 \leq x < m) \vee (\forall y(m \leq y \leq x \rightarrow A(y)))\}$$

Then B is internal, by the internal definition principle, and contains $\mathcal{N}$. If B does not contain an infinite natural number ω, then $B = \mathcal{N}$, contradicting Lemma 2.11(iii). Hence, there exists an ω with the required properties.

(ii). If the set ${}^*\mathcal{N} \setminus A$ is nonempty, then, being internal, it must contain a least element $\omega > 0$. This implies that $A(\omega - 1)$, hence, by hypothesis, $A(\omega)$ holds, which is a contradiction.

□

Although Robinson's construction of enlargements is sufficient to carry out many basic arguments in analysis, for example see Lemma 2.31, we require a slightly stronger property to develop the theory in Section 3;

Definition 2.13. *(i). We say that an enlargement* ${}^*\mathcal{R}$ *of* ${}^*\mathcal{R}$ *is* $\aleph_1$-*saturated, if given a countable decreasing sequence of internal sets* $(A_m)_{m\in\mathcal{N}}$, *with each* $A_m \neq \emptyset$, *then* $\bigcap_{m\in\mathcal{N}} A_m \neq \emptyset$.

(ii). We say that an enlargement ${}^*\mathcal{R}$ *of* $\mathcal{R}$ *has countable comprehension, if given a sequence* $(A_m)_{m\in\mathcal{N}}$ *of internal sets, with each* $A_m \subset A$, *and* A *internal, then there exists an internal sequence* $(A_m)_{m\in{}^*\mathcal{N}}$, *with each* $A_m \subset A$, *extending it, ([3]).*

Lemma 2.14. *The definitions* (i) *and* (ii) *are equivalent. Moreover, there exists an enlargement* ${}^*\mathcal{R}$ *satisfying* (i), *therefore* (ii).

[3]In the paper [24], Loeb says that ${}^*\mathcal{R}$ is a denumerably comprehensive enlargement, if ,given a standard set S, a sequence $(A_m)_{m\in\mathcal{N}}$ of internal sets, with each $A_m \in {}^*S$, the sequence $(A_m)_{m\in\mathcal{N}}$ is the restriction to $\mathcal{N}$ of an internal function $f : {}^*\mathcal{N} \rightarrow {}^*S$. This property is actually equivalent to (ii), even though we require that S is standard. The proof is an easy exercise, but is essentially contained in Lemma 2.14.

Proof. (i) implies (ii). Let;

$B_m = \{f : f\ is\ a\ function\ with\ domain\ {}^*\mathcal{N}\ and\ range\ {}^*(V_\tau(\mathcal{R})),$ $such\ that\ \bigwedge_{1\leq i\leq m} f(i) = a_i \wedge \forall i({}^*\mathcal{N}(i) \rightarrow f(i) \subset a)\}$

where $\{a_1, \ldots, a_m, a\}$ are names for the sets $\{A_1, \ldots, A_m, A\}$ and τ is the type of A. Then B_m is internal by the internal definition principle. Moreover, one can formulate a sentence in $\mathcal{R}$, which says that for any given set W of type (τ) and cardinality m (this can be done using only quantifiers), there exists a function $f : \mathcal{N} \rightarrow W$ enumerating it, hence, by the transfer principle, if W is an internal set of type (τ) and cardinality m, there exists an internal function $f : {}^*\mathcal{N} \rightarrow W$ enumerating it. Therefore, $B_m \neq \emptyset$. Clearly, the sequence $(B_m)_{m\in\mathcal{N}}$ is decreasing, so, using (i), we can find an f with the required properties belonging to $\bigcap_{m\in\mathcal{N}} B_m$.

(ii) implies (i). Let $(A_m)_{m\in\mathcal{N}}$ be a given countable decreasing sequence of internal sets, with $A_m \neq \emptyset$. Using type considerations, we can assume that each $A_m \subset {}^*S$, with S standard, and each $A_m \in {}^*T$, with T standard. By (ii), there exists an internal function $f : {}^*\mathcal{N} \rightarrow T$, with $f(i) = A_i$, for $i \in \mathcal{N}$. Let;

$$B = \{n \in {}^*\mathcal{N} : \exists g \forall (i \leq n)(g(i) = f(i) \wedge g(i) \neq \emptyset \wedge g(n) \subseteq g(i))\}$$

Then B contains $\mathcal{N}$ as the condition is satisfied, for all finite n by f. By overflow, there exists an infinite ω and g such that $g(\omega) \subseteq g(i)$, for all $i \in \mathcal{N}$, and $g(\omega) \neq \emptyset$. Hence, as $g(i) = A_i$, for all $i \in \mathcal{N}$, $\bigcap_{m\in\mathcal{N}} A_m \neq \emptyset$ as required.

To prove the existence of an enlargement, satisfying (i), you need to show that condition (i) is equivalent to the requirement that all 1-types over countable subsets of ${}^*(V(\mathcal{R}))$ are satisfied, $(*)$. To obtain $(*)$, you can probably proceed as in the first order context, see [11], adapting this argument to higher order structures. The basic idea is to realise all the 1-types over countable subsets of a given model $\mathcal{M}$, using compactness and transfinite induction. Suppose that you have constructed a chain of models $\{\mathcal{M}_i : i < \omega\}$, with $\mathcal{M}_0 = M$ and ω an ordinal. If ω is not a limit ordinal, use compactness to realise all the 1-types over countable subsets of $\mathcal{M}_{\omega-1}$ inside a model $\mathcal{M}_{\omega-1} \subset \mathcal{M}_\omega$, if ω is a limit ordinal, take a union $\bigcup_{i<\omega} \mathcal{M}_i$. You then satisfy $(*)$ in the first order context, when $Card(\omega) \geq \aleph_1$. Due to the technical language involved in Robinson's method, it would take some time to write this down rigorously, however, I do not envisage any problems. In [12], it is claimed that $\aleph_1$-saturation follows from the Mostowski construction, however, I haven't seen a proof of this result.

□

Before turning to some simple applications of nonstandard analysis, we mention two more related logical notions, which will be indispensable throughout this thesis.

Definition 2.15. *Let A in ${}^*\mathcal{R}$ be an internal set. We say that A is $*$-finite or hyperfinite, if there exists an internal $f : [0,\omega] \rightarrow A$, enumerating the elements of A, where $\omega \in {}^*\mathcal{N}$ and $[0,\omega] = \{x \in {}^*\mathcal{N} : 0 \leq x \leq \omega\}$. We define the internal cardinality $Card(A) = \omega$.*

Remarks 2.16. *Observe that this definition makes sense, as if two functions f and f' enumerate A, defined on $[0,\omega]$ and $[0,\omega']$ respectively, then $\omega = \omega'$. This follows immediately from the transfer principle, and the obvious fact that any finite set in $\mathcal{R}$ has a uniquely defined cardinality. The important point to note, here, is that the set of finite sets in $\mathcal{R}$ of type τ is definable as;*

$$\{x \in V_\tau(\mathcal{R}) : \exists n \exists f(\forall y(x(y) \leftrightarrow \exists i(0 \leq i \leq n \wedge f(i) = y)))\}$$

so we can transfer facts about finite sets in $\mathcal{R}$ to $*$-finite sets in ${}^*\mathcal{R}$.

Lemma 2.17. *If A and B are $*$-finite, then $A \cap B$ and $A \cup B$ are $*$-finite. If C is internal and $C \subseteq A$, then C is $*$-finite.*

Proof. Immediate, using the transfer principle, and elementary facts about finite sets. □

More generally, we note the following;

Lemma 2.18. *Let A be an internal set, and let $SF(A) = \{B : B \subseteq A, B \text{ is internal and } hyperfinite\}$. Then $SF(A)$ is internal, and there exists an internal function $Card : SF(A) \rightarrow {}^*\mathcal{N}$, which assigns to each element of $SF(A)$ its internal cardinality. If A is $*$-finite, then $SF(A)$ is $*$-finite.*

Proof. As in the previous lemma. □

Following from this, we have;

Definition 2.19. *Let A be $*$-finite of type τ, and let $f : A \rightarrow {}^*\mathcal{R}$ be internal. Let $\Gamma \subset V_\tau(\mathcal{R}) \times V_{(\tau,0)}(\mathcal{R}) = \{(C, g) : C \in Finite(V_\tau(\mathcal{R})), g \in Function(C, \mathcal{R})\}$, where $Finite(V_\tau(\mathcal{R}))$ consists of the sets in $V_\tau(\mathcal{R})$ which are finite, and $Function(C, \mathcal{R})$ consists of the functions $g : C \rightarrow \mathcal{R}$. Then we define the hyperfinite or $*$-finite sum;*

$${}^*\sum_{x \in A} f(x)$$

to be ${}^*S(A, f)$ where $S : \Gamma \rightarrow \mathcal{R}$ is the standard function assigning a standard pair (C, g), the finite sum;

$$\sum_{x \in C} g(x)$$

and ${}^*S$ is the transfer of the function S to ${}^*\mathcal{R}$

Remarks 2.20. *Note this is a good definition as the transfer ${}^*\Gamma$ consists exactly of pairs (A, f), where A is $*$-finite of type τ, and $f : A \rightarrow {}^*\mathcal{R}$ is internal.*

We now turn to some simple applications of Robinson's construction, which we will use later in the thesis;

Theorem 2.21. *Let $(s_n)_{n \in \mathcal{N}}$ be a standard infinite sequence, then the following are equivalent;*

(i). $(s_n)_{n \in \mathcal{N}}$ is bounded in $\mathcal{R}$.

(ii). All the elements of $(s_n)_{n \in {}^*\mathcal{N}}$ are finite in ${}^*\mathcal{R}$.

(iii). $(s_n)_{n \in {}^*\mathcal{N}}$ is bounded by a finite number in ${}^*\mathcal{R}$.

Proof. (i) implies (iii). Choose $m \in \mathcal{R}$ with $|s_n| < m$, for all $n \in \mathcal{N}$. Let f be the standard function enumerating the sequence $(s_n)_{n \in \mathcal{N}}$ and let σ be the sentence;

$$\forall x \forall y \forall z (f(x) = y \wedge z = |y| \rightarrow z < m)$$

Then σ is true in $\mathcal{R}$, hence, true in ${}^*\mathcal{R}$, in particular the transferred function ${}^*f$, enumerating $(s_n)_{n \in \mathcal{N}}$, is bounded by m.

(iii) implies (ii). Obvious.

(ii) implies (i). Let r be any infinite positive number in ${}^*\mathcal{R}$, and let σ be the sentence;

$\forall x \forall y \forall z (f(x) = y \wedge z = |y| \rightarrow z < r)$;

Then σ is true in ${}^*\mathcal{R}$. Let σ' be the sentence;

$\exists w \forall x \forall y \forall z (f(x) = y \wedge z = |y| \rightarrow z < w)$;

Then σ' is true in ${}^*\mathcal{R}$, hence is true in $\mathcal{R}$. Therefore, there exists an $m \in \mathcal{R}$ bounding the sequence $(s_n)_{n \in \mathcal{N}}$.

□

Theorem 2.22. *Let $(s_n)_{n \in \mathcal{N}}$ be a standard infinite sequence, then the following are equivalent;*

(i). $lim_{n \rightarrow \infty} s_n = s$.

(ii). $s_n \simeq s$ *for all infinite* n.

Proof. (i) implies (ii). Let $\epsilon > 0$, and choose $m \in \mathcal{N}$ with $|f(n) - s| < \epsilon$ for all $n \in \mathcal{N}$ with $n > m$. Let σ be the sentence;

$\forall x ((N(x) \wedge x > m) \rightarrow |f(x) - s| < \epsilon)$ $(*)$

Then σ is true in $\mathcal{R}$, and, hence, also true in ${}^*\mathcal{R}$. Let ω be any infinite natural number, then $\omega > m$, hence $|{}^*f(\omega) - s| < \epsilon$. As ϵ was arbitrary, we conclude that $s_\omega \simeq s$ as required.

(ii) implies (i). Let $\epsilon > 0$, and let ω be any infinite natural number, then $s_\omega \simeq s$, hence, $|{}^*f(\omega) - s| < \epsilon$. It follows that $(*)$ holds in ${}^*\mathcal{R}$, for any given infinite natural number ω', replacing m. Let σ' be the sentence;

$\exists w \forall x ((N(x) \wedge x > w) \rightarrow |f(x) - s| < \epsilon)$

Then σ' holds in ${}^*\mathcal{R}$, hence, in $\mathcal{R}$. This produces an m, with $|f(n) - s| < \epsilon$, for all $n \in \mathcal{N}$ with $n > m$. As ϵ was arbitrary, (i) is shown.

□

The following is a slight generalisation of Theorem 3.5.13 of [36];

Theorem 2.23. *Let $(s_n)_{n\in{}^*\mathcal{N}}$ be an internal sequence, enumerated by an internal g, not necessarily the transfer of a standard one. Suppose there exists an infinite $\omega \in {}^*\mathcal{N}$ with $s_{\omega'} \simeq 0$, for all infinite ω' with $\omega' < \omega$, then;*

$lim_{n\rightarrow\infty}(^\circ s_n) = 0$ *in the standard sense.*

Proof. Let $\epsilon > 0$ be standard, then $|g(\omega')| < \epsilon$ for all infinite ω' with $\omega' < \omega$. Let;

$A = \{m \in {}^*\mathcal{N} : |g(n)| < \epsilon, if\ m \leq n < \omega\}$

Then A is internal and contains arbitrarily small positive infinite numbers. By the underflow principle, see Lemma 2.12, footnote 2, it contains a positive finite number $m_0 \in \mathcal{N}$. In particular, $|g(n)| < \epsilon$, for all $n \in \mathcal{N}$, with $n \geq m_0$. It follows that $|^\circ(g(n))| \leq \epsilon$, for all $n \in \mathcal{N}$, with $n \geq m_0$. Hence, as ϵ was arbitrary, the result follows. □

Theorem 2.24. *Let $f : \mathcal{R} \rightarrow \mathcal{R}$ be a standard function, and $b \in \mathcal{R}$, then the following are equivalent;*

(i). $lim_{x\rightarrow b}f(x) = c$.

(ii). ${}^*f(x) \simeq c$ *for all* $x \in \mu(b) \setminus \{b\}$.

Proof. (i) implies (ii). Let $\epsilon > 0$ be standard, then there exists a standard $\delta > 0$ such that $|f(x) - c| < \epsilon$, for all $x \in \mathcal{R}$ with $0 < |x - b| < \delta$. By transfer, the same holds in ${}^*\mathcal{R}$ for ${}^*f$. In particular, if $x \in \mu(b) \setminus \{b\}$, then, $0 < |x - b| < \delta$, so $|{}^*f(x) - c| < \epsilon$. As ϵ was arbitrary, we must have ${}^*f(x) \simeq c$ as required.

(ii) implies (i). Let δ be a positive infinitesimal, let $\epsilon > 0$ be standard, and let σ be the sentence;

$\forall x(0 < |x - b| < \delta \rightarrow |f(x) - c| < \epsilon)$

Then σ is true in ${}^*\mathcal{R}$. Let σ' be the sentence;

$\exists w \forall x (0 < |x - b| < w \rightarrow |f(x) - c| < \epsilon)$

Then σ' is true in ${}^*\mathcal{R}$, hence, true in $\mathcal{R}$. As ϵ was arbitrary, it follows that $lim_{x \rightarrow b} f(x) = c$, as required.

□

Theorem 2.25. *Let $f : \mathcal{R} \rightarrow \mathcal{R}$ be a standard function, and $b \in \mathcal{R}$, then the following are equivalent;*

(i). f is continuous at b.

(ii). ${}^*f(x) \simeq {}^*f(b)$ for all $x \in \mu(b)$, ([4]).

Proof. Immediate from the definition of continuity and Theorem 2.24.

□

Remarks 2.26. *There is a remarkable analogue of Theorem 2.25 in the context of a complete Zariski structure C, due to Zilber, see* [45]. *This states that if $W \subset C^n \times C^m$ is a closed finite cover of C^n, then, given $(b, c) \in W$, if $b' \in \mathcal{V}_b$, there exists $c' \in \mathcal{V}_c$ with $(b', c') \in W$, where $\mathcal{V}_b$ and $\mathcal{V}_c$ are the infinitesimal neighborhoods of b and c. For a continuous function $f : C \rightarrow C$ with respect to the topology on C, these conditions are satisfied, taking $W = graph(f)$. Hence, by Remarks 2.2 and the GAGA principle, we obtain Theorem 2.25, (i) implies (ii), in the case of the projective line, $P^1({}^*\mathcal{C})$ over the hyper complex numbers. It would be interesting to find more connections between Robinson's and Zilber's work, see also the remarks on topology at the end of this section.*

A simple consequence of Theorem 2.24 is the following. Of course, (ii) is how Newton originally defined the derivative, in [26], although a working calculation involving this definition first appears in [29], see also [33]. Robinson's following theorem makes this idea rigorous.

Theorem 2.27. *Let $f : \mathcal{R} \rightarrow \mathcal{R}$ be a standard function, and $b \in \mathcal{R}$, then the following are equivalent;*

(i). $f'(b) = c$ where f' denotes the derivative of f.

(ii). $\frac{{}^*f(x) - {}^*f(b)}{x - b} \simeq c$ for all $x \in \mu(b) \setminus \{b\}$.

[4] This is often applied in the following form; if $x \simeq y$ belong to ${}^*\mathcal{R}_{fin}$, then ${}^*f(x) \simeq {}^*f(y)$. This follows from the facts that x and y have a standard part in $\mathcal{R}$, and $\simeq$ is an equivalence relation.

Proof. Immediate, by Theorem 2.24, and the standard definition of the derivative.

□

Robinson also gives a non standard account of integration and a non-standard proof of the Fundamental Theorem of Calculus. The non-standard definition of integration is the one Newton intuitively uses in [27] and [28]. However, Newton's proof of the Fundamental Theorem of Calculus, in [28], differs from Robinson's, and is interesting to study in its own right. In [33], I found a rigorous justification of Newton's proof, using standard arguments, it would be interesting, however, to do this using only non-standard methods.

Robinson's account of integration requires that the standard function f should be continuous on the standard interval $[a,b]$. We make the following definitions;

Definition 2.28. *Let $\eta \in {}^*\mathcal{N} \setminus \mathcal{N}$. Let $A = \{i \in {}^*\mathcal{N} : 0 \leq i < \eta\}$ and $B = \{x \in {}^*\mathcal{R} : a \leq x \leq b\}$. By the infinitesimal partition π_η of $[a,b]$, I mean the internal function $\pi_\eta : A \rightarrow B$, defined by;*

$$\pi_\eta(i) = a + \tfrac{i}{\eta}(b-a)$$

Definition 2.29. *Let g be internal, defined on B, then we set $\int_{\eta,a,b} g$ to be the $*$-finite sum;*

$$\tfrac{1}{\eta}{}^*\textstyle\sum_{i\in A} g(\pi_\eta(i))$$

Remarks 2.30. *Observe that this is a good definition, by Remarks 2.10, and the fact that $Im(\pi_\eta)$ is $*$-finite.*

We now claim;

Lemma 2.31. *Let the standard f be continuous on the standard interval $[a,b]$, then;*

$$^\circ(\textstyle\int_{\eta,a,b} {}^*f) = \int_a^b f dx \ (*)$$

In particular, the definition of the non-standard integral, as the left hand side of $(*)$, is independent of the choice of $\eta \in {}^*\mathcal{N} \setminus \mathcal{N}$.

Proof. Let $\epsilon > 0$ be standard, then, by Darboux's theorem, there exists a $\delta > 0$, such that;

$$|C - \sum_{i=0}^{m-1} f(x_i)(x_{i+1} - x_i)| < \epsilon$$

for all $\{x_0, \ldots, x_m\}$ with $x_0 = a < x_1 < \ldots < x_m = b$, and $|x_{i+1} - x_i| < \delta$, for $i = 0, \ldots, m-1$, where $C = \int_a^b f dx$. In particular, taking $M > \frac{b-a}{\delta}$, we have that;

$$|C - \frac{1}{n}\sum_{i=0}^{n-1} f(a + \frac{i}{n}(b-a))| < \epsilon \text{ for } n \geq M.$$

As M is standard and, therefore $\eta > M$, we have, by the transfer principle and Definition 2.19;

$$|C - \frac{1}{\eta}{}^*\sum_{i=0}^{\eta-1} {}^*f(\pi_\eta(i))| < \epsilon$$

Hence, by Definition 2.29;

$$|C - \int_{\eta,a,b} {}^*f| < \epsilon$$

As ϵ was arbitrary, the result follows.

□

Remarks 2.32. *We have slightly modified the definition of the non-standard integral, given in* [36]*, in order to emphasise the connection with hyperfinite sums, and internal integrals, which we will discuss in the next section. However, the basic idea is the same. The use of a hyperfinite partition occurs in each of the following sections, so it is crucial for the reader to understand this last argument. One should observe that it depends on the fact that f is continuous. Without this assumption, that is assuming only that f is Lebesgue integrable on* $[a, b]$*, the definition, given above, is not correct. Consider the standard function defined on* $[0, 1]$*, taking the value* 0 *on* $\mathcal{Q}$ *and* 1 *on* $\mathcal{Q}^c$*. Then, the Lebesgue of integral of f is* 1*, but, as is easily checked, the non-standard integral, as defined above, is* 0*. Robinson attempted to overcome this difficulty by defining a non-standard Lebesgue integral on a bounded interval, however, his definition is quite strange, and he doesn't show that it coincides with the standard Lebesgue integral. A more intuitive definition using samples is given in* [8]*. They prove that there exists a sample, (which is essentially a partition with varying step size), representing Lebesgue measure, and use this result, to show that in the* *case of a bounded Lebesgue measurable function f on a bounded interval, their definition coincides with the Lebesgue integral. This result was later extended in [7], to include unbounded Lebesgue measurable functions on unbounded intervals. The definition is still slightly unsatisfactory, however, due to the varying step size of a sample, In [41], a regular sample, that is one with constant step size, is found representing Lebsegue measure on a bounded interval, the proof uses the ergodic theorem. I am not sure whether, this result has been shown over the real line, and whether it can be used to give a nonstandard definition of the Lebesgue integral for unbounded functions. One can also give a definition of the Lebesgue integral using the theory of lifts from Loeb integration. This is the approach taken in Section 7. We will consider Loeb integration in the following section. The main advantage of this method is that one can take a constant step size in the integral. The disadvantage is that the lift, though preserving the integral, is* not *the transfer of the original function, so it may change other properties of the function concerned. In Section 5, the use of lifts is avoided, because the restriction that f belongs to the Schwarz class, allows a simple hyperfinite definition of the Lebesgue integral, as above, to go through.*

Robinson also gives a nonstandard proof of the Fundamental Theorem of Calculus, but I will not include it here. I finish this section by giving a brief account of nonstandard topology, which is used in Sections 6 and 7.

Let $T = (A, \Omega)$ be a topological space, where Ω is a set of type $((0))$, and whose elements are the open sets of A. Using Robinson's theory, we can construct an enlargement ${}^*T = ({}^*A, {}^*\Omega)$ of T. ${}^*T$ is not, in general, a topological space, in the standard sense, as its open sets ${}^*\Omega$ are only preserved by taking internal unions, rather than arbitrary unions.

Definition 2.33. *Let $p \in A$, then we define the monad $\mu(p)$ to be;*

$\bigcap_{U_\nu \in \Omega_p} {}^*U_\nu$

where $\Omega_p \subset \Omega$ consists of all the standard open sets containing p.

If $r \in {}^*A$, then we say that r is near standard if there exists $p \in A$, with $r \in \mu(p)$.

We have the following fundamental result;

Theorem 2.34. *(i). T is Hausdorff iff for all distinct $p, q \in T$, $\mu(p) \cap \mu(q) = \emptyset$.*

(ii). T is compact iff all points of ${}^*A$ are near standard.

In particular, if T is Hausdorff and compact, there exists a unique standard part mapping $st : {}^*A \rightarrow A$.

Proof. The proof can be found in [36], Theorems 4.1.8 and 4.1.13. □

Definition 2.35. *Let T be compact, and let $f : {}^*T \rightarrow {}^*\mathcal{R}$ be internal and finite, in the sense of Definition 2.1, then we say that f is S-continuous if, for all $x, y \in {}^*T$, if $x \simeq y$, then $f(x) \simeq f(y)$.*

Theorem 2.36. *Let T be compact and let $f : {}^*T \rightarrow {}^*\mathcal{R}$ be S-continuous, then $g : T \rightarrow \mathcal{R}$, defined by $g(x) = {}^\circ f(x)$ is continuous.*

Proof. The proof can be found in [36], Theorems 4.5.8 and 4.5.10. □

3. Loeb Integration and Measure

Many of the results in this section are based on [24], these results were later extended, but, at the time of writing I could only find them stated in [12], and so proved them myself. I later found some analagous results in [1].

We work in an $\aleph_1$-saturated enlargement, ${}^*\mathcal{M}$, where $\mathcal{M}$ contains $\mathcal{R}$. The reader can assume that $\mathcal{M} = \mathcal{R}$, but, it is not necessary for the proofs to go through. We assume that X is an internal set in $\mathcal{M}$, and that $\mathfrak{A}$ is an internal collection of internal subsets of X. We require that $\mathfrak{A}$ is an algebra on X, see [44]. We observe the following;

Lemma 3.1. *Let $(X, \mathfrak{A})$ be as above. Then $*$-finite unions of sets in $\mathfrak{A}$ are in $\mathfrak{A}$.*

Proof. We define;

$$I = \{n \in {}^*\mathcal{N} : \forall f([\forall i(0 \leq i \leq n \wedge f(i) \in \mathfrak{A})] \rightarrow [\exists y(y \in \mathfrak{A} \wedge \forall z(X(z) \rightarrow (y(z) \leftrightarrow \exists i(0 \leq i \leq n \wedge z \in f(i)))))])\}$$

I consists of those n for which all cardinality $n+1$ $*$-finite unions of sets in $\mathfrak{A}$, belong to $\mathfrak{A}$. I is internal, moreover, $I(0)$ obviously holds, and if $I(n)$ holds then $I(n+1)$ holds, as $\mathfrak{A}$ is an algebra. By the internal induction principle, $I = {}^*\mathcal{N}$, so the claim follows.

□

Remarks 3.2. *It does* not *necessarily follow that internal unions of sets in* $\mathfrak{A}$*, indexed by* ${}^*\mathcal{N}$ *belong to* $\mathfrak{A}$*. Algebras* $\mathfrak{A}$ *with this stronger property are called* ${}^*\sigma$*-algebras. If we assume that* X *itself is* $*$*-finite, so* $\mathfrak{A}$ *is* $*$*-finite, then,* $\mathfrak{A}$ *is a* ${}^*\sigma$*-algebra. This follows immediately from the transfer principle and the fact that finite algebras, that is algebras with finitely many elements, are* σ*-algebras. In the applications from Sections 5,6 and 7, we will always have that* X *is* $*$*-finite, so the reader can assume this stronger property.*

However, it does not follow that $\mathfrak{A}$ is a σ-algebra, see also the preceding remark. This follows from;

Lemma 3.3. *Given* $A_n \in \mathfrak{A}$*, with* $n \in \mathcal{N}$*, if* $A_0 \subset \bigcup_{n=1}^{\infty} A_n$*, then there exists an* $m \in \mathcal{N}$*, with* $A_0 \subset \bigcup_{n=1}^{m} A_n$*. In particular, countable unions of disjoint nonempty sets in* $\mathfrak{A}$ *are* not *in* $\mathfrak{A}$*.*

Proof. By countable comprehension, we can find an internal sequence $(A_n)_{n\in^*\mathcal{N}}$ extending the sequence $(A_n)_{n\in\mathcal{N}}$, with $A_n \in \mathfrak{A}$, for $n \in^* \mathcal{N}$. The set $B = \{r \in^* \mathcal{N} : A_0 \subseteq \bigcup_{n=1}^{r} A_n\}$ is internal, nonempty and contains all infinite $\omega \in^* \mathcal{N}$, hence, by underspill, it contains a finite element m. This shows the first claim. For the second part, suppose that there exist $A_n \in \mathfrak{A}$, for $n \geq 1$, nonempty and disjoint, such that $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{A}$. Then, letting $A_0 = \bigcup_{n=1}^{\infty} A_n$, we can apply the first part, to find an $m \in \mathcal{N}$, with $A_0 \subset \bigcup_{n=1}^{m} A_n$, contradicting the fact that the sets A_n are disjoint. □

We now suppose that there exists a finitely additive, hence $*$-finitely additive, internal function $\nu : \mathfrak{A} \rightarrow^* \mathcal{R}_{\geq 0}$. (Again, ν will be ${}^*\sigma$ additive if X is $*$-finite). For $A \in \mathfrak{A}$, we set $\mu(A) =^\circ \nu(A)$, so $\mu : \mathfrak{A} \rightarrow \mathcal{R}_{\geq 0} \cup +\infty$ is finitely additive. Let $\mathfrak{M}$ denote the σ-algebra on X generated by $\mathfrak{A}$. Then we have the following theorem due to Loeb, in [24];

Theorem 3.4. *Let* $(X, \mathfrak{A})$*,* $\mathfrak{A} \subset \mathfrak{M}$*, and* ν*,*μ *be as above. Then;*

(i). μ *has a standard* σ*-additive extension to* $\mathfrak{M}$*, which we also denote by* μ*, such that, if* $B \in \mathfrak{M}$*, then* $\mu(B) = inf_{A\in\mathfrak{A}, B\subset A}\mu(A)$*.*

(ii). If $\mu(X) < \infty$*, then, the extension is unique, and, for* $B \in \mathfrak{M}$*,* $\mu(B) = sup_{A\in\mathfrak{A}, B\supset A}\mu(A)$*, moreover, there exists* $A \in \mathfrak{A}$*, with* $\mu(A \Delta B) = 0$*, where* Δ *denotes symmetric difference.*

Proof. The proof is in [24], but some parts need clarification.

(i). Observe that μ is countably additive on $\mathfrak{A}$, by Lemma 3.3 and the fact that μ is finitely additive. Hence, we can construct an extension by Caratheodory's Theorem, see [44]. In this construction, you define an outer measure λ on the σ-algebra $\mathfrak{G}$ of all subsets of X, by defining $\lambda(G) = inf(\sum_{n\in\mathcal{N}}\mu(A_n))$, where $A_n \in \mathfrak{A}$ and $G \subseteq \bigcup_{n\in\mathcal{N}} A_n$, and restricting to $\mathfrak{M}$, $(*)$. Now suppose that $B \in \mathfrak{M}$, with $\mu(B) < \infty$, and $\epsilon > 0$. Then, using $(*)$, we can find an increasing sequence $(A_n)_{n\in\mathcal{N}}$, with $B \subset C = \bigcup_{n\in\mathcal{N}} A_n$, and $\mu(C) < \mu(B) + \epsilon$. Now extend $(A_n)_{n\in\mathcal{N}}$ to an internal sequence $(A_n)_{n\in{}^*\mathcal{N}}$, using countable comprehension. By overflow, we can find an infinite ω, such that $A_n \subset A_{n+1}$, for all n, with $0 \leq n \leq \omega$, and with $\nu(A_\omega) < \mu(B) + \epsilon$. Clearly $B \subset C \subset A_\omega$, and $\mu(A_\omega) < \mu(B) + \epsilon$, as $\mu(A_\omega) \simeq \nu(A_\omega)$. As ϵ was arbitrary, the result follows, when $\mu(B) < \infty$. The case $\mu(B) = \infty$ is obvious.

(ii). If $\mu(X) < \infty$, the uniqueness result follows from Caratheodory's Theorem. Now, if $B \in \mathfrak{M}$, then, using (i) and $\mu(X) < \infty$, $\mu(B) = \mu(X) - \mu(B^c) = \mu(X) - inf_{A\in\mathfrak{A}, B^c\subset A}\mu(A) = \mu(X) - (\mu(X) - sup_{A\in\mathfrak{A}, B\supset A}\mu(A)) = sup_{A\in\mathfrak{A}, B\supset A}\mu(A)$, showing the second claim. Using this result, we can find an increasing sequence $(A_n)_{n\in\mathcal{N}_{>0}}$, belonging to $\mathfrak{A}$, with $A_n \subset B$, and $\mu(B - A_n) < \frac{1}{n}$, $(*)$. Using countable comprehension and overflow, we can extend this to an increasing sequence $(A_n)_{n\in\mathcal{N}_{0<n\leq\omega}}$, in $\mathfrak{A}$, for some infinite ω, with $\mu(B) - \nu(A_n) < \frac{1}{n}$ (this is an internal condition), for $0 < n \leq \omega$. In particular, it follows that $\mu(B) - \nu(A_\omega) < \frac{1}{\omega}$, hence $\mu(B) - \mu(A_\omega) = 0$. Let $A = \bigcup_{n\in\mathcal{N}_{>0}} A_n$. Then $A \subset B$ and $A \subset A_\omega$, $(**)$, moreover, $\mu(B - A) = 0$ by $(*)$. Therefore, $\mu(B) = \mu(A) = \mu(A_\omega)$, so using $(**)$, we have $\mu(A_\omega \Delta B) = 0$ as required.

□

Remarks 3.5. *It was later shown by Ward Henson, that even if $\mu(X) = \infty$, one obtains a unique extension to $\mathfrak{M}$, but I haven't been able to find a reference.*

Definition 3.6. *Let $f : X \rightarrow {}^*\mathcal{R}$ be internal, then we say that f is $\mathfrak{A}$-measurable, if, for all $a \in {}^*\mathcal{R}$;*

$\{x \in X : f(x) < a\}$ *and* $\{x \in X : f(x) \leq a\}$ *belong to* $\mathfrak{A}$.

Remarks 3.7. *Note that if $\mathfrak{A}$ is a ${}^*\sigma$-algebra, then only one condition is necessary, as, for example;*

$\{x \in X : f(x) \leq a\}$

$= \bigcap_{n\in {}^*\mathcal{N}}\{x \in X : f(x) < a + \frac{1}{n}\}$

$= (\bigcup_{n\in {}^*\mathcal{N}}(\{x \in X : f(x) < a + \frac{1}{n}\})^c)^c$

Here, the first equality follows using the fact that ${}^*\mathcal{Q}$ is dense in ${}^*\mathcal{R}$. Now use Remarks 3.2.

Observe also, that if $\mathfrak{A}$ is a ${}^*\sigma$-algebra, then, if f and g are $\mathfrak{A}$-measurable, so is $f + g$ and fg. This follows from the corresponding fact for σ-algebras, see [37]*. Otherwise, it needn't be true. However, it is always true that f measurable implies $|f|$ is measurable, without the ${}^*\sigma$ assumption. We also have, in the case that $\mathfrak{A}$ is a ${}^*\sigma$-algebra, that measurability, as defined above, is equivalent to either of the conditions $f^{-1} : {}^*\tau \rightarrow \mathfrak{A}$, $f^{-1} : {}^*\mathfrak{D} \rightarrow \mathfrak{A}$, where ${}^*\tau$ and ${}^*\mathfrak{D}$ are the transfers of the usual topology τ and the Borel field $\mathfrak{D}$ on $\mathcal{R}$ to ${}^*\mathcal{R}$. This follows immediately from the corresponding fact for σ-algebras, and measurable functions taking values in $\mathcal{R}$. In Section 5, we replace ${}^*\mathcal{R}$ by ${}^*\mathcal{C}$, the transfer of the complex numbers $\mathcal{C}$. As we work there always with ${}^*\sigma$-algebras, the easiest way to define measurability is to require that $f^{-1} : {}^*\tau \rightarrow \mathfrak{A}$, where ${}^*\tau$ is the transfer of the complex topology. This is equivalent to $Re(f)$ and $Im(f)$ being measurable in the sense defined above.*

We have the following theorem;

Theorem 3.8. *If $f : X \rightarrow {}^*\mathcal{R}$ is $\mathfrak{A}$-measurable, then ${}^\circ f : X \rightarrow \mathcal{R} \cup \{+\infty, -\infty\}$ is $\mathfrak{M}$-measurable.*

Proof. Just observe that;

$\{x \in X :^\circ f(x) < a\} = \bigcup_{n=1}^{\infty}\{x \in X : f(x) < a - \frac{1}{n}\}$

and use the fact that $\mathfrak{M}$ is a σ-algebra. □

We now turn to Loeb integration theory. We first require the important concept of an internal integral. We already came across this idea in Definition 2.29.

Definition 3.9. *Let X be $*$-finite and let $f : X \rightarrow {}^*\mathcal{R}$ be $\mathfrak{A}$-measurable. Then, for $A \in \mathfrak{A}$, we define the internal integral;*

$\int_A f d\nu = {}^*G(A, f, X, \mathfrak{A}, \nu)$

where ${}^*G$ is the transfer of the standard function G, defined by setting;

$G(A_0, f_0, X_0, \mathfrak{A}_0, \nu_0) = \int_{A_0} f_0 d\nu_0$

for a standard tuple $(A_0, f_0, X_0, \mathfrak{A}_0, \nu_0)$, consisting of a finite measure space $(X_0, \mathfrak{A}_0, \nu_0)$, a finite subset $A_0 \subset A$, and a measurable function $f_0 : A \rightarrow \mathcal{R}$. See [37], *Definition 1.23 and 1.31, for a definition of integration on measure spaces, in terms of simple functions.*

Remarks 3.10. *Observe that for a finite measure space, the integral is always defined, hence, the internal integral always takes a value in ${}^*\mathcal{R}$. Moreover, if $\mathfrak{A}_0$ consists of all the subsets of X_0, the standard integral is just $\sum_{x\in A_0} f_0(x)\nu_0(x)$, hence the internal integral becomes a hyperfinite sum. If ν is a counting measure on $\mathfrak{A}$, defined by $\nu(A) = \frac{Card(A)}{Card(X)}$, for an internal subset $A \subset X$, we obtain the hyperfinite sum $\frac{1}{\eta}{}^*\sum_{x\in A} f(x)$, as in Definition 2.29, where $\eta = Card(X)$. More specifically, if X is a hyperfinite interval, see Sections 4,5,6 and 7, that is $X = \{x \in {}^*\mathcal{R} : \frac{a}{\eta} \leq x < \frac{b}{\eta}\}$, where $a, b \in {}^*\mathcal{Z}$, $\mathfrak{A}$ is the set of all internal unions of intervals of the form $[\frac{j}{\eta}, \frac{j+1}{\eta})$, where $a \leq j < b$, $j \in {}^*\mathcal{Z}$ and ν is the counting measure given by $\nu([\frac{j}{\eta}, \frac{j+1}{\eta})) = \frac{1}{\eta}$, then the internal integral takes the form;*

$\int_X f d\nu = \frac{1}{\eta}\sum_{j=a}^{b-1} f(\frac{j}{\eta})$

This is an extremely important notion, as it is used repeatedly, in all of the following sections. Some authors, see [12], *prefer to use a discrete version of the hyperfinite interval, in which $X = \{\frac{j}{\eta} : a \leq j < b, j \in {}^*\mathcal{Z}\}$, $\mathfrak{A}$ is the set of internal subsets, and ν is the counting measure given by $\nu(x) = \frac{1}{\eta}$, for $x \in X$. Of course the two interpretations are equivalent and the internal integral takes the same form. In fact, this is the approach taken in Section 6.*

Observe also that as finite measure spaces are σ-algebras, we obtain transferred versions of standard results such as the dominated convergence theorem, see [37], *for the internal integral, however, these properties are not very useful.*

If X is not $*$-finite, we can still transfer G, now defined on spaces $(X_0, \mathfrak{A}_0, \nu_0)$, where $\mathfrak{A}_0$ is an algebra, ν_0 is a finitely additive measure, with $\nu_0(X_0) < \infty$, and f_0 is measurable with respect to $\mathfrak{A}_0$, by using the definition in [37]. *However, there is a slight technical problem in that G may be undefined for certain functions. We can easily overcome this difficulty, by restricting to the transfer of integrable functions, and defining the internal integral to be $+\infty$ otherwise, (we then set $^\circ(+\infty) = +\infty$). Of course, this problem does not arise, if X is $*$-finite or f is bounded by a value in $\mathcal{R}$. Observe that as $\mathfrak{A}_0$ is not necessarily a σ-algebra, we do not obtain the usual limit theorems(transferred) for the internal integral, although Rudin's definition still works.*

We observe the following, which will also be required in Section 6;

Lemma 3.11. *Suppose $f : X \rightarrow {}^*\mathcal{R}_{fin}$ is internal, then f is bounded by some $n \in \mathcal{N}$.*

Proof. Suppose f is unbounded, and let;

$$B = \{n \in^* \mathcal{N} : \exists x(X(x) \wedge |f(x)| > n)\}$$

Then B is internal and contains $\mathcal{N}$. By overspill, it contains some infinite $\omega \in {}^*\mathcal{N}$, contradicting the hypothesis that f takes values in ${}^*\mathcal{R}_{fin}$.

□

We now have the following, in [24], which I include for the convenience of the reader;

Theorem 3.12. *Assume $\mu(X) < \infty$, and $f : X \rightarrow {}^*[-n, n]$ is $\mathfrak{A}$-measurable, with $n \in \mathcal{N}$, then for $A \in \mathfrak{A}$;*

$$^\circ \int_A f d\nu = \int_A {}^\circ f d\mu$$

where $\int_A f d\nu$ is the internal integral on $(X, \mathfrak{A}, \nu)$ and $\int_A {}^\circ f d\mu$ is the standard integral, on $(X, \mathfrak{M}, \mu)$.

Proof. By considering the function $f\chi_A$, instead of f, it is clearly sufficient to prove the result when $A = X$. We can also assume that for some $\delta > 0$ in $\mathcal{R}$, $f(x) \geq \delta$, for all $x \in X$, $(*)$. For, if $k = sup_{x\in X}|^\circ f(x)| + 2\delta$, then, if we show that $\int_X (f + k)d\nu \simeq \int_X {}^\circ(f + k)d\mu$, then we clearly obtain $\int_X f d\nu \simeq \int_X {}^\circ f d\mu$, as $k\nu(X) \simeq k\mu(x)$. Let

$D=\{r\in\mathcal{R}:\mu((^{\circ}f)^{-1}[r])>0\}$. Then D is finite or countably infinite in $\mathcal{R}$, $(**)$. This follows, as for $n\in\mathcal{N}_{>0}$, the set $D_n=\{r\in\mathcal{R}:\mu((^{\circ}f)^{-1}[r])>\frac{1}{n}\}$ must be finite, as $\mu(X)<\infty$, and $D=\bigcup_{n\in\mathcal{N}_{>0}}D_n$. Now let $M=\mu(X)+1$, and $\epsilon\in\mathcal{R}_{>0}$. By considering the intervals, $\{(0,\frac{\epsilon}{2M}),(\frac{\epsilon}{2M},\frac{2\epsilon}{2M}),\ldots\}$, and using $(*)$,$(**)$, we can find $m\in\mathcal{N}$, and $\{y_0,\ldots,y_m\}$, with $0=y_0<y_1<\ldots<y_m$, such that $y_i\notin D$, for $0\leq i\leq m$, and $y_{i+1}-y_i<\frac{\epsilon}{2M}$, for $0\leq i\leq m-1$. Let;

$$\underline{S_\nu}=\sum_{i=1}^m y_{i-1}\nu(f^{-1}[y_{i-1},y_i))$$

$$\overline{S_\nu}=\sum_{i=1}^m y_i\nu(f^{-1}[y_{i-1},y_i))$$

$$\underline{S_\mu}=\sum_{i=1}^m y_{i-1}\mu(^{\circ}f^{-1}[y_{i-1},y_i))$$

$$\overline{S_\mu}=\sum_{i=1}^m y_i\mu(^{\circ}f^{-1}[y_{i-1},y_i))$$

Then, by the definition of the standard integral, and the transferred definition of the internal integral, we have;

$$\underline{S_\nu}\leq\int_X fd\nu\leq\overline{S_\nu}$$

$$\underline{S_\mu}\leq\int_X {}^{\circ}fd\mu\leq\overline{S_\mu}\ (***)$$

Clearly, we have that;

$$\overline{S_\nu}-\underline{S_\nu}\leq\frac{\epsilon}{2M}\sum_{i=1}^m\nu(f^{-1}[y_{i-1},y_i))<\frac{\epsilon}{2}$$

and, similarly, $\overline{S_\mu}-\underline{S_\mu}<\frac{\epsilon}{2}$, $(****)$. Now observe that;

$$^{\circ}f^{-1}(y_{i-1},y_i)\subseteq f^{-1}(y_{i-1},y_i)\subseteq f^{-1}[y_{i-1},y_i]\subseteq{}^{\circ}f^{-1}[y_{i-1},y_i]\ (\dagger)$$

and;

$$\mu(^{\circ}f^{-1}[y_{i-1},y_i])=\mu(^{\circ}f^{-1}[y_{i-1},y_i))=\mu(^{\circ}f^{-1}(y_{i-1},y_i))\ (\dagger\dagger)$$

as $y_{i-1}\notin D$ and $y_i\notin D$.

Moreover;

$$\mu(^{\circ}f^{-1}(y_{i-1},y_i))\leq\mu(f^{-1}(y_{i-1},y_i))\simeq\nu(f^{-1}(y_{i-1},y_i))\leq\nu(f^{-1}[y_{i-1},y_i))$$

$\leq \nu(f^{-1}[y_{i-1}, y_i]) \simeq \mu(f^{-1}[y_{i-1}, y_i]) \leq \mu(^\circ f^{-1}[y_{i-1}, y_i])$ (†††)

by definition of ν and μ, and (†). It follows that all the quantities from (††) and (†††) lie in the same monad, as the same quantity lies at both ends. Therefore, $\underline{S_\nu} \simeq \underline{S_\mu}$ and $\overline{S_\nu} \simeq \overline{S_\mu}$. It follows from this and $(***)$, $(****)$, that $|\int_X f d\nu - \int_X {^\circ f} d\mu| < \epsilon$. As ϵ was arbitrary, the result follows. □

The following theorem, in [24], requires quite a bit of explanation not given in the paper;

Theorem 3.13. *Assume $\mu(X) < \infty$, and $g : X \rightarrow \mathcal{R} \cup \{+\infty, -\infty\}$ is $\mathfrak{M}$-measurable, then there exists an $f : X \rightarrow {^*\mathcal{R}}$, which is $\mathfrak{A}$-measurable, such that $^\circ f = g$ almost everywhere, with respect to μ.*

Proof. We first reduce to the case that g is bounded by some $M \in \mathcal{N}$. Assume that g is unbounded, and let $C_n = g^{-1}([-n, n])$, $W_{+\infty} = g^{-1}(+\infty)$, $W_{-\infty} = g^{-1}(-\infty)$ and $B_n = C_n \cup W_{+\infty} \cup W_{-\infty}$, for $n \in \mathcal{N}_{>0}$. Then, $B_n \subseteq B_{n+1}$ and $X = \bigcup_{n \in \mathcal{N}_{>0}} B_n$, $(*)$. Using Theorem 3.4(ii), we can find a sequence $(A_i)_{i \in \mathcal{N}_{>0}}$, with the properties that the A_i are disjoint, belong to $\mathfrak{A}$, $A_i \subset B_i$ and $\mu(B_n - \bigcup_{1 \leq i \leq n} A_i) < \frac{1}{n}$, $(**)$. Moreover, as $\{C_n, W_{+\infty}, W_{-\infty}\}$ are disjoint, we can assume that $A_{i,+\infty} = A_i \cap W_{+\infty}$, $A_{i,-\infty} = A_i \cap W_{-\infty}$ and $A'_i = A_i \cap C_i$ belongs to $\mathfrak{A}$, for each i. Now consider the function $g'_i = g\chi_{A'_i}$. Using the result for bounded functions, proved below, there exists an $\mathfrak{A}$-measurable function f'_i with $^\circ f'_i = g'_i$ a.e $d\mu$. Without loss of generality, we can suppose f'_i is supported on A'_i, by replacing f'_i with $f'_i\chi_{A'_i}$, see remark below. Now let $f_i = f'_i + \omega\chi_{A_{i,+\infty}} - \omega\chi_{A_{i,-\infty}}$, where ω is infinite, then clearly $^\circ f_i = g_i$, a.e $d\mu$, where $g_i = g\chi_{A_i}$, $(***)$. Moreover, f_i is still $\mathfrak{A}$-measurable, again see remark below, and supported on A_i. Now, by countable comprehension, we can extend the sequences $(A_i)_{i \in \mathcal{N}_{>0}}$, $(f_i)_{i \in \mathcal{N}_{>0}}$ to $(A_i)_{i \in {^*\mathcal{N}_{>0}}}$, $(f_i)_{i \in {^*\mathcal{N}_{>0}}}$. By overflow, we can find an infinite $\eta \in {^*\mathcal{N}_{>0}}$, such that the A_i are disjoint and the f_i are $\mathfrak{A}$-measurable, for $0 < i \leq \eta$. Using the fact that the initial segment $(0, \eta]$ of ${^*\mathcal{N}_{>0}}$ is $*$-finite, we can form the $*$-finite sum $f =^* \sum_{0 < i \leq \eta} f_i$. By transfer, and using the fact that the A_i are disjoint, f is $\mathfrak{A}$-measurable, see following remark. We claim that $^\circ f = g$ a.e $d\mu$, $(****)$. Let $A = \bigcup_{i \in \mathcal{N}} A_i$, then, using $(***)$ and the fact that the A_i are disjoint, we have $^\circ f|A = g|A$, a.e $d\mu$. Moreover, using $(*)$,$(**)$;

$$\mu(X - A) = \mu(X) - \mu(A) = lim_{n \rightarrow \infty} \mu(B_n - \bigcup_{1 \leq i \leq n} A_i) = 0$$

which clearly shows $(****)$.

For the bounded case, we can, without loss of generality, assume that $0 \leq g \leq M$. By dividing the range of g into finitely many intervals of length $\frac{1}{n}$, we can find an increasing sequence of *simple* $\mathfrak{M}$-measurable functions $\{g_n\}_{n\in\mathcal{N}_{>0}}$, with $0 \leq g - g_n \leq \frac{1}{n}$, $(\dagger)$. By Theorem 3.4(ii)(last part), there exists a sequence of $\mathfrak{A}$-measurable functions $\{f_n\}_{n\in\mathcal{N}_{>0}}$, taking values in ${}^*[0, M]$ with ${}^\circ f_n = g_n$ a.e $d\mu$, $(\dagger\dagger)$. Then, using Remarks 3.7 and Theorem 3.12, for $n < m$;

$$^\circ \int_X |f_n - f_m| d\nu = \int_X |{}^\circ f_n - {}^\circ f_m| d\mu = \int_X |g_n - g_m| d\mu \leq \frac{\mu(X)}{n} \ (\dagger\dagger\dagger)$$

As usual, we can extend the sequence $\{f_n\}_{n\in\mathcal{N}_{>0}}$ to a sequence of $\mathfrak{A}$-measurable functions $(f_n)_{n\in{}^*\mathcal{N}_{0<n\leq\omega'}}$, for some infinite ω', such that each f_n takes values in ${}^*[0, M]$. Now, for $n \in \mathcal{N}_{>0}$, let;

$$C_n = \{m \in (n, \omega'] : \int_X |f_n - f_m| d\nu < \tfrac{\mu(X)+1}{n}\}.$$

Then C_n is internal, and, by $(\dagger\dagger\dagger)$, $D_n = \bigcap_{0<i\leq n} C_n \neq \emptyset$. Hence, using $\aleph_1$-saturation, $\bigcap_{n\in\mathcal{N}_{>0}} C_n \neq \emptyset$. Therefore, we can find an infinite ω, with $\omega \leq \omega'$, such that $\int_X |f_n - f_\omega| d\nu < \frac{\mu(X)+1}{n}$, for every $n \in \mathcal{N}_{>0}$, $(\dagger\dagger\dagger\dagger)$. We now have, using Theorem 3.12,$(\dagger)$,$(\dagger\dagger)$,$(\dagger\dagger\dagger\dagger)$, for $n \in \mathcal{N}_{>0}$;

$$\int_X |g -^\circ f_\omega| d\mu \leq \int_X |g -^\circ f_n| d\mu + \int_X |{}^\circ f_n -^\circ f_\omega| d\mu$$

$$\simeq \int_X |g - g_n| d\mu + \int_X |f_n - f_\omega| d\nu$$

$$< \frac{\mu(X)}{n} + \frac{\mu(X)+1}{n} < \frac{2}{n}(\mu(X) + 1)$$

As $n \in \mathcal{N}_{>0}$ was arbitrary, $g =^\circ f_\omega$ a.e $d\mu$, as required.

□

Remarks 3.14. *In the previous proof, we made some measurability claims. These follow immediately, if we assume that* $\mathfrak{A}$ *is* $*$*-finite, by Remarks 3.7. Otherwise, we can use the following principle;*

Suppose A *and* B *are disjoint subsets of* $\mathfrak{A}$*. If* f_A *and* f_B *are* $\mathfrak{A}$*-measurable functions, supported on* A *and* B *respectively, then* $f_A + f_B$ *is* $\mathfrak{A}$*-measurable. If* f *is* $\mathfrak{A}$*-measurable, then* $f\chi_A$ *is* $\mathfrak{A}$*-measurable.*

The proof is an easy exercise.

We now turn to various generalisations of Loeb's results. The proofs are my own, although I later found some analagous results in [1]. The reader should be aware, however, that these results assume $\mathfrak{A}$ is a ${}^*\sigma$-algebra. We replace the measure space $(X,\mathfrak{M},\mu)$ by its completion, which I denote by $(X,\mathfrak{M}_L,\mu_L)$. This space is referred to as the Loeb space associated to $(X.\mathfrak{A},\nu)$. We observe the following;

Lemma 3.15. *The results of Theorem 3.4, Theorem 3.8, Theorem 3.12, Theorem 3.13 hold with $(X,\mathfrak{M}_L,\mu_L)$ replacing $(X,\mathfrak{M},\mu)$.*

Proof. Theorem 3.4 follows immediately from the definition of a completion, see [37]. Theorem 3.8 is obvious. Theorem 3.12 follows from the fact that, if g is $\mathfrak{M}$-integrable, then;

$$\int_B g d\mu = \int_B g d\mu_L$$

for any $B \in \mathfrak{M}_L$. In order to see this, first check the result for $\mathfrak{M}$-measurable simple functions, $(*)$. Then, without loss of generality, assume $g \geq 0$. g can be written as in increasing limit of simple $\mathfrak{M}$-measurable functions, see [37], Theorem 1.17. Now apply the Monotone Convergence Theorem, see [37], Theorem 1.26, and $(*)$, to obtain the result. Finally, Theorem 3.13 follows from the fact that if h is $\mathfrak{M}_L$-measurable, then we can find g which is $\mathfrak{M}$-measurable, such that $g = h$ a.e $d\mu$, see [37], Theorem 8.12 (Lemma 1). □

Theorem 3.16. *Let $f : X \rightarrow^* \mathcal{R}_{\geq 0}$, be $\mathfrak{A}$-measurable, then we always have that;*

$$\int_A {}^\circ f d\mu_L \leq^\circ \int_A f d\nu$$

where we allow $+\infty$ on either side. We do not *assume that $\mu_L(X) < \infty$.*

Proof. We can clearly assume that $A = X$.

Case 1. $\mu_L(X) < \infty$.

Let $C_n = \{x \in X : f(x) > n\}$ and let $B = \bigcap_{n\in\mathcal{N}} C_n$, then we can assume that $\mu_L(B) = 0$, $(*)$. If not, using Lemma 3.15 (3.4(ii)), we can find $D \subset B$, with $D \in \mathfrak{A}$, such that $\nu(D) = r > 0$, where r is standard. By overflow, there exists an infinite ω, with $f \geq \omega$ on D. Then $\int_D {}^\circ f d\mu_L = \infty$, see [37], and ${}^\circ \int_D f d\nu \geq^\circ (\omega r) = \infty$, as $r > 0$, using transfer. Clearly then, the result follows. Assuming $(*)$, let $f_n = min\{f, n\}$, for $n \in \mathcal{N}$. We clearly have, by transfer, that;

$\int_X f_n d\nu \leq \int_X f d\nu$, for $n \in \mathcal{N}$

Hence;

$^\circ \int_X f d\nu \geq lim_{n\rightarrow\infty} {}^\circ \int_X f_n d\nu$ $(**)$

Observe that $^\circ f_n = min(^\circ f, n)$, so, using $(*)$ and Lemma 3.15 (3.8), $\{^\circ f_n\}$ is an increasing sequence of $\mathfrak{M}_L$-measurable functions, converging to $^\circ f$ a.e dμ_L. By the monotone convergence theorem, and Lemma 3.15 (3.12)

$\int_X {}^\circ f d\mu_L = lim_{n\rightarrow\infty} \int_X {}^\circ f_n d\mu_L = lim_{n\rightarrow\infty} {}^\circ \int_X f_n d\nu$

Hence, the result follows from $(**)$.

Case 2. $\mu_L(X) = \infty$.

We can assume that $^\circ f$ vanishes outside a set B of σ-finite measure with respect to μ_L, $(*)$. For suppose not, and $^\circ f > 0$ on a set B which is not σ-finite. Let $C_n = \{x \in X : {}^\circ f > \frac{1}{n}\}$, for $n \in \mathcal{N}_{>0}$, then C_n is increasing and $B = \bigcup_{n\in\mathcal{N}_{>0}} C_n$. As B is not σ-finite, we can assume that $\mu_L(C_n) = \infty$, for some n, $(**)$. As $^\circ f \geq 0$, this implies that $\int_X {}^\circ f d\mu_L = \infty$. Let $A_n = \{x \in X : f > \frac{1}{n}\}$, for $n \in \mathcal{N}_{>0}$. Then A_n is an increasing sequence of $\mathfrak{A}$-measurable sets, and clearly $A_n = B_n$. Therefore, we can find $n \in \mathcal{N}$, with $\nu(A_n) = \omega$, for some infinite ω, by $(**)$. By transfer, and the fact that $f \geq 0$, $\int_X f d\nu \geq \frac{\omega}{n}$, so $^\circ \int_X f d\nu = \infty$, as n is standard. This shows the result. Now assuming $(*)$, suppose that $B = \bigcup_{n\in\mathcal{N}} D_n$, with $\mu_L(D_n) < \infty$, and $\{D_n\}$ increasing. By Lemma 3.15 (3.4(i)), we can choose an increasing sequence $\{A_n\}$ of $\mathfrak{A}$-measurable sets, such that $D_n \subset A_n$, $\mu_L(A_n) < \infty$, and $B \subset A$, where $A = \bigcup_{n\in\mathcal{N}} A_n$. Then $^\circ f$ vanishes outside A. Hence, using the MCT, Case 1, and previous observations;

$\int_X {}^\circ f d\mu_L = lim_{n\rightarrow\infty} \int_X {}^\circ f \chi_{A_n} d\mu_L \leq lim_{n\rightarrow\infty} {}^\circ \int_X f \chi_{A_n} d\nu \leq {}^\circ \int_X f d\nu$

as required.

□

From Definition 3.17 to Remarks 3.21, we will assume that $\nu(X)$ is finite, or equivalently $\mu_L(X)<\infty$.

Definition 3.17. *Let $f:X\rightarrow^{*}\mathcal{R}$ be $\mathfrak{A}$-measurable, with $\nu(X)$ finite. Then f is S-integrable, if;*

(i). $\int_X|f|d\nu$ is finite.

(ii). If $A\in\mathfrak{A}$, with $\nu(A)\simeq 0$, then $\int_A|f|d\nu\simeq 0$.

Theorem 3.18. (a). *f is S-integrable iff f^+ and f^- are S-integrable.*

(b). *If f is $\mathfrak{A}$-measurable, then f is S-integrable iff $|f|$ is S-integrable.*

Proof. (a). Clearly f is $\mathfrak{A}$-measurable iff f^+ and f^- are $\mathfrak{A}$-measurable. If f is S-integrable, then as $\int_A|f^+|d\nu\leq\int_A|f|d\nu$, for any $A\in\mathfrak{A}$, we have that f^+ is S-integrable, and similarly for f^-. If f^+ and f^- are S-integrable, then, $|f|$ is $\mathfrak{A}$- measurable, and, for $A\in\mathfrak{A}$;

$$\int_A|f|d\nu=\int_{A\cap f\geq 0}|f^+|d\nu+\int_{A\cap f\leq 0}|f^-|d\nu$$

so clearly, f is S-integrable.

(b). If f is S-integrable, then $|f|$ is $\mathfrak{A}$-measurable, so $|f|$ is S-integrable by the definition. If f is $\mathfrak{A}$-measurable and $|f|$ is S-integrable, then again f is S-integrable, by the definition.

□

Lemma 3.19. *Let $f:X\rightarrow^{*}\mathcal{R}$ be $\mathfrak{A}$-measurable, with $\nu(X)$ finite. Then;*

f is S-integrable iff $\int_{|f|>K}|f|d\nu\simeq 0$ for all infinite K

Proof. Left to right. Suppose that f is S-integrable, and let K be infinite. Clearly, $A=\{x\in X:|f|(x)>K\}$ belongs to $\mathfrak{A}$. Hence, by Definition 3.17(ii), it is sufficient to prove that $\nu(A)\simeq 0$, $(*)$. We have, using elementary properties of the internal integral, and Definition 3.17(i);

$$K\nu(A)\leq\int_A|f|d\nu\leq\int_X|f|d\nu=c$$

where c is finite. Therefore, as K is infinite, $\nu(A)\simeq 0$, showing $(*)$.

Right to left. We first show Definition 3.17(i). Suppose that $\int_X|f|d\nu$ is not finite, then we can find an infinite $\omega\in^*\mathcal{N}_{>0}$ with $2\omega<\int_X|f|d\nu$. As $\nu(X)$ is finite, there exists an infinite $\eta\in^*\mathcal{N}_{>0}$, with $\eta\nu(X)<\omega$. Let $A=\{x\in X:|f|(x)>\eta\}$, then A is $\mathfrak{A}$-measurable, and;

$$2\omega<\int_X|f|d\nu=\int_A|f|d\nu+\int_{A^c}|f|d\nu\leq\epsilon+\eta\nu(X)<1+\omega$$

where ϵ is an infinitesimal. This implies that $\omega<1$, which is a contradiction. We now show Definition 3.17(ii). We first claim the following;

$$^\circ\int_A|f|d\nu=lim_{n\rightarrow\infty}{}^\circ\int_A min(|f|,n)d\nu \text{ for } A\in\mathfrak{A},\ (\dagger)$$

Suppose $(\dagger)$ fails, then we can find $r<s$ in $\mathcal{R}$, such that;

$$\int_A min(|f|,n)d\nu<r<s<\int_A|f|d\nu$$

holds for all $n\in\mathcal{N}$. By overflow, there exists an infinite $K>0$, such that $\int_A min(|f|,K)d\nu<r$, $(\dagger\dagger)$. Let $A'=\{x\in A:|f|(x)>K\}$. Then, by the assumption in the lemma, the definition of A' and $(\dagger\dagger)$;

$$\int_A|f|d\nu=\int_{A'}|f|d\nu+\int_{A\setminus A'}|f|d\nu\simeq\int_{A\setminus A'}|f|d\nu\leq\int_A min(|f|,K)d\nu$$

$$<r<s<\int_A|f|d\nu\ (\dagger\dagger\dagger)$$

Taking standard parts in $(\dagger\dagger\dagger)$, we obtain a contradiction. Thus we can assume $(\dagger)$. Now fix a standard $\epsilon>0$, and choose $n\in\mathcal{N}$, such that $\int_X|f|d\nu-\int_X min(|f|,n)d\nu<\epsilon$, $(\dagger\dagger\dagger\dagger)$. Define $F:X\rightarrow^*\mathcal{R}$ by setting;

$F(x)=f(x)$, if $|f(x)|\leq n$
$F(x)=-n$, if $f(x)<-n$
$F(x)=n$, if $f(x)>n$.

Then F and $f-F$ are $\mathfrak{A}$-measurable, and $|f-F|=|f|-|F|$, so $\int_X|f-F|d\nu<\epsilon$, by $(\dagger\dagger\dagger\dagger)$. We now have that, for $A\in\mathfrak{A}$;

$$\int_A|f|d\nu\leq\int_A|f-F|d\nu+\int_A Fd\nu\leq\epsilon+n\nu(A)$$

Therefore, if $\nu(A) \simeq 0$, then $\int_A |f| d\nu < 2\epsilon$. As ϵ was arbitrary, we have that $\int_A |f| d\nu \simeq 0$.

This proves the claim.

□

Theorem 3.20. *Let $f : X \rightarrow^* \mathcal{R}$ be $\mathfrak{A}$-measurable, with $\mu_L(X)$ finite. Then the following are equivalent;*

(i). f is S-integrable.

(ii). $^\circ f$ is integrable with respect to μ_L, and;

$^\circ \int_A f d\nu = \int_A {}^\circ f d\mu_L$ *for any $A \in \mathfrak{A}$.*

Proof. We can assume, using Theorem 3.18, that $f \geq 0$.

(i) implies (ii). By Theorem 3.16, and Definition 3.17(i), $^\circ f$ is integrable with respect to μ_L, hence, the standard sequence $min(^\circ f, n)_{n \in \mathcal{N}}$ converges a.e dμ_L to $^\circ f$, $(*)$. Therefore, using (†) from Lemma 3.19, Lemma 3.15 (3.12), and $(*)$;

$$^\circ \int_A f d\nu = lim_{n \rightarrow \infty} {}^\circ \int_A min(f, n) d\nu = lim_{n \rightarrow \infty} \int_A min(^\circ f, n) d\mu_L$$

$$= \int_A {}^\circ f d\mu_L$$

for any $A \in \mathfrak{A}$, as required.

(ii) implies (i). By Lemma 3.19, it is sufficient to prove that for $K > 0$ infinite, $\int_{f > K} f d\nu \simeq 0$, $(**)$. Let $A = \{x \in X : f(x) > K\}$, then $\mu_L(A) = 0$, as $^\circ f|A = \infty$, and $^\circ f$ is integrable with respect to μ_L. Therefore, $\int_A^\circ f d\mu_L = 0$. If $(**)$ fails, $\int_A f d\nu \geq \epsilon > 0$, where ϵ is standard. Therefore, $^\circ \int_A f d\nu > \int_A {}^\circ f d\mu_L$, which contradicts the assumption of (ii).

□

Remarks 3.21. *One can actually prove the implication from (ii) to (i) in the previous theorem, with a weaker assumption. Namely;*

Assume $f : X \rightarrow^* \mathcal{R}_{\geq 0}$ *is* $\mathfrak{A}$*-measurable, with* $\mu_L(X)$ *finite, then, if* ${}^\circ f$ *is integrable with respect to* μ_L*, and;*

$${}^\circ \int_X f d\nu = \int_X {}^\circ f d\mu_L \ (*).$$

then f *is* S*-integrable.*

To see this, just follow through the proof of (ii) *implies* (i) *above, with the same notation. To obtain the final contradiction of the assumption* $(*)$*, observe, using Theorem 3.16, that;*

$${}^\circ \int_X f d\nu = {}^\circ \int_A f d\nu + {}^\circ \int_{A^c} f d\nu \gneq \int_{A^c} {}^\circ f d\mu_L = \int_X {}^\circ f d\mu_L$$

From Definition 3.22 to Remarks 3.25, we assume that $\mu_L(X) = \infty$.

Definition 3.22. *Let* $f : X \rightarrow^* \mathcal{R}$ *be* $\mathfrak{A}$*-measurable, with* $\nu(X)$ *infinite. Then* f *is* S*-integrable, if;*

Conditions $(i), (ii)$ *of Definition 3.17 hold* and*;*

(iii). If $A \in \mathfrak{A}$ *and* $f \simeq 0$ *on* A*, then* $\int_A |f| d\nu \simeq 0$.

Remarks 3.23. *Observe that the extension of Theorem 3.18 to this case still holds. However, Lemma 3.19 fails;*

If f *is* S*-integrable in the sense of Definition 3.21, then, for all infinite* K*;*

$$\int_{|f|>K} |f| d\nu \simeq 0$$

but the converse is not true.

To see this, one can check that the same proof, for one direction, as in Lemma 3.19 works. A simple counterexample to the other direction is as follows;

Let $X = [0, \omega] \subset^* \mathcal{N}$ *be a hyperfinite interval, with* ω *infinite,* $\mathfrak{A}$ *the algebra of internal subsets, and counting measure* ν*. Let* $f|[0, \omega] = \frac{1}{\omega}$*, then clearly,* $f \simeq 0$ *on* X*, but* $\int_X f d\nu = \frac{\omega}{\omega} = 1$*, contradicting Definition 3.22(iii).*

Theorem 3.24. *Let $f : X \rightarrow^* \mathcal{R}$ be $\mathfrak{A}$-measurable, with $\mu_L(X)$ infinite. Then the following are equivalent;*

$(i)'$. f is S-integrable.

$(ii)'$. ${}^\circ f$ is integrable with respect to μ_L, and;

${}^\circ \int_A f d\nu = \int_A {}^\circ f d\mu_L$ *for any* $A \in \mathfrak{A}$.

Proof. Again, using the beginning of Remarks 3.23, we can assume that $f \geq 0$.

$(i)'$ implies $(ii)'$. Suppose f is S-integrable. Then, by Theorem 3.16, we have, as before, that ${}^\circ f$ is integrable with respect to μ_L. As in Theorem 3.16, Case 2, and using Definition 3.22(i), we can assume that ${}^\circ f$ vanishes outside a σ-finite set B. Following the proof of Theorem 3.16, Case 2, we can find an increasing sequence $\{E_n\}$ of $\mathfrak{A}$-measurable sets, such that $\mu_L(E_n) < \infty$ and $B \subset E$, where $E = \bigcup_{n \in \mathcal{N}} E_n$. Then ${}^\circ f$ vanishes outside E. Now, using the MCT, and the fact that $f\chi_{E_n}$ is S-integrable, we can apply Theorem 3.20, to obtain;

$\int_A {}^\circ f d\mu_L = lim_{n\rightarrow\infty} \int_A {}^\circ f \chi_{E_n} d\mu_L = lim_{n\rightarrow\infty} {}^\circ \int_A f\chi_{E_n} d\nu$ $(*)$

Now fix $\epsilon > 0$ standard, then, by $(*)$, we can choose n such that;

$0 < c - \int_A f\chi_{E_m} d\nu < \epsilon$, for $m \geq n$, where $c = \int_A {}^\circ f d\mu_L$ $(**)$

Now using the fact that $(**)$ is an internal condition, we can apply countable comprehension and overflow, to obtain an infinite $\omega > 0$, and a $\mathfrak{A}$-measurable E_ω with $B \subset E_\omega$, such that;

$0 < \int_A {}^\circ f d\mu_L - \int_A f\chi_{E_\omega} d\nu < \epsilon$ $(***)$

As $B \subset E_\omega$, ${}^\circ f = 0$ on E_ω^c and $f \simeq 0$ on E_ω^c. Therefore, using Definition 3.22(iii), $\int_A f d\nu \simeq \int_{A\cap E_\omega} f d\nu$. It follows, by $(***)$, that;

$0 < \int_A {}^\circ f d\mu_L - \int_A f d\nu < \epsilon$

As ϵ was arbitrary, we obtain that ${}^\circ \int_A f d\nu = \int_A {}^\circ f d\mu_L$, as required.

$(ii)'$ implies $(i)'$. We have to check conditions $(i), (ii), (iii)$ from Definition 3.22.

Condition (i). By the assumption in $(ii)'$, we clearly have that $\int_X f d\nu$ is finite.

Condition (ii). We first claim that for all infinite $K > 0$, $\int_{f>K} f d\nu \simeq 0$, $(\sharp)$. This follows exactly as in the second part of Theorem 3.20. Now, using $(\sharp)$, you can show that;

$^\circ \int_A |f| d\nu = lim_{n\rightarrow\infty} {}^\circ \int_A min(|f|, n) d\nu$ for $A \in \mathfrak{A}$, $(\sharp\sharp)$

exactly as in the proof of Lemma 3.19. Finally, use the same truncation argument, as in Lemma 3.19, to obtain the result that, for $A \in \mathfrak{A}$, with $\nu(A) \simeq 0$, $\int_A f d\nu \simeq 0$, as required.

Condition (iii). Suppose that $A \in \mathfrak{A}$, with $f|A \simeq 0$. and, for contradiction, that $\int_A f d\nu > r > 0$, where r is standard, $(\sharp\sharp\sharp)$. We have that $^\circ f|A = 0$, so $\int_A^\circ f d\mu_L = 0$. Then, using $(\sharp\sharp\sharp)$ and Theorem 3.16, we must have that;

$\int_X f d\nu > r + \int_{A^c} f d\nu \geq r + \int_{A^c} {}^\circ f d\mu_L > 0 + \int_{A^c} {}^\circ f d\mu_L = \int_X {}^\circ f d\mu_L$

contradicting the assumption in $(ii)'$.

$\square$

Remarks 3.25. *Again, as in Remarks 3.21, one can actually prove the implication from (ii) to (i) in the previous theorem, with a similar weaker assumption;*

Assume $f : X \rightarrow^* \mathcal{R}_{\geq 0}$ is $\mathfrak{A}$-measurable, with $\mu_L(X)$ infinite, then, if $^\circ f$ is integrable with respect to μ_L, and;

$^\circ \int_X f d\nu = \int_X {}^\circ f d\mu_L$ $(*)$.

then f is S-integrable.

Again, follow through the proof of $(ii)'$ implies $(i)'$ above. Checking (i) is the same. In (ii), to obtain $(\sharp)$, you need to use the same trick as in Remarks 3.21. Then the rest of the proof goes through. In (iii), the proof is the same.

For the convenience of the reader, we combine Definition 3.17 and Definition 3.22, no assumption on the measure is made;

Definition 3.26. *Let $f : X \rightarrow^{*} \mathcal{R}$ be $\mathfrak{A}$-measurable. Then f is S-integrable, if;*

(i). $\int_X |f| d\nu$ is finite.

(ii). If $A \in \mathfrak{A}$, with $\nu(A) \simeq 0$, then $\int_A |f| d\nu \simeq 0$.

(iii). If $A \in \mathfrak{A}$ and $f \simeq 0$ on A, then $\int_A |f| d\nu \simeq 0$.

Remarks 3.27. *Observe that if $\nu(X)$ is finite, and f is S-integrable in the sense of Definition 3.17, then f is S-integrable in the sense of Definition 3.26. This follows because, if $f \simeq 0$ on A, then for any standard $\epsilon > 0$, $\int_A |f| d\nu \leq \epsilon\nu(X)$. Clearly, this implies that $\int_A |f| d\nu \simeq 0$.*

Using this remark, we then have the following convenient portmanteau version of Theorem 3.16, Theorem 3.20, Remarks 3.21, Theorem 3.24 and Remarks 3.25. The proof is clear.

Theorem 3.28. *Suppose $f : X \rightarrow^{*} \mathcal{R}$ is $\mathfrak{A}$-measurable, then the following are equivalent;*

(i). f is S-integrable in the sense of Definition 3.26.

(ii). ${}^{\circ}f$ is μ_L-integrable and ${}^{\circ}\int_X |f| d\nu = \int_X |{}^{\circ}f| d\mu_L$.

(iii). ${}^{\circ}f$ is μ_L-integrable and ${}^{\circ}\int_X |f| d\nu \leq \int_X |{}^{\circ}f| d\mu_L$.

Definition 3.29. *Anderson*

Let $f : X \rightarrow {}^{*}\mathcal{R}$ be $\mathfrak{A}$-measurable. Then we say f is finite if;

(i). There exists an $n \in \mathcal{N}$, with $|f(x)| < n$, for all $x \in X$.

(ii). f is supported on a set A with $\nu(A)$ finite.

The following result is required in Section 5.

Theorem 3.30. *Anderson's Criteria*

Let $f : X \rightarrow^{*} \mathcal{R}$ *be* $\mathfrak{A}$*-measurable.*

(i). *If* F *is* S*-integrable, with* $|f| \leq F$*, then* f *is* S*-integrable.*

(ii). *If* $\mathfrak{A}$ *is a* ${}^{*}\sigma$*-algebra, then* f *is* S*-integrable iff there exists a sequence of finite functions* $(f_n)_{n \in \mathcal{N}}$ *such that;*

${}^{\circ}(\int_X |f - f_n| d\nu) \rightarrow 0$ *as* $n \rightarrow \infty$.

Proof. (i). This requires one direction of (ii), we leave the details to the reader, or see [1].

(ii). We will just prove one direction. Suppose there exists a sequence $(f_n)_{n \in \mathcal{N}}$, with the given properties. Using Theorem 3.18(i), and redefining $f_n' = f_n \chi_{f \geq 0}$ and $f_n'' = f_n \chi_{f \leq 0}$, we can assume that $f \geq 0$. We can also assume that $0 \leq f_n \leq f$, by redefining $f_n' = mid(0, f_n, f)$. We can assume that f_n is increasing, by taking $f_n' = max\{f_1, \ldots, f_n\}$. Finally, we can suppose that $sup\{f_n\} < n$, by taking $f_n' = min\{n - 1, f_n\}$. We check conditions $(i), (ii), (iii)$ of Definition 3.26.

(i). We have that;

$\int_X f d\nu = \int_X (f - f_n) d\nu + \int_X f_n d\nu$

$< 1 + n\nu(\{x : f_n(x) \neq 0\}$

for n sufficiently large, this is finite by Definition 3.29(ii).

(ii). Let $\epsilon > 0$ be standard, and choose n such that ${}^{\circ} \int_X (f - f_n) d\nu < \frac{\epsilon}{2}$, $(*)$. If $A \in \mathfrak{A}$, with $\nu(A) < \frac{\epsilon}{2n}$, then;

$\int_A f d\nu \leq n\nu(A) + \int_A (f - f_n) d\nu < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$

Therefore, if $\nu(A) \simeq 0$, $\int_A f d\nu < \epsilon$, for all $\epsilon > 0$, so $\int_A f d\nu \simeq 0$.

(iii). If $A \in \mathfrak{A}$, with $f(A) \subset \mu(0)$, then, using Lemma 2.12(i), there exists $\eta \simeq 0$ with $0 \leq f|_A < \eta$. Then, choosing n as in $(*)$ above;

$\int_A f d\nu < \frac{1}{\eta}\nu(\{x : f_n(x) \neq 0\} + \frac{\epsilon}{2} < \frac{\epsilon}{2}$

As ϵ was arbitrary, $\int_A f d\nu \simeq 0$.

□

We now generalise Theorem 3.13;

Theorem 3.31. *Lifting Theorem*

Let $g : X \to \mathcal{R}$ be $\mathfrak{M}_L$-measurable, then g is integrable with respect to μ_L iff there exists an S-integrable $f : X \to^* \mathcal{R}$, such that $^\circ f = g$ a.e $d\mu_L$.

Proof. One direction is obvious, if g has an S-integrable lifting f, then g is μ_L integrable, by Theorem 3.28(ii).

For the other direction, we can assume that $g \geq 0$. As usual, we divide the proof into two cases.

Case 1. $\mu_L(X) < \infty$. If g is bounded, then by the proof of Lemma 3.15(Theorem 3.13), we can find a bounded f, such that $^\circ f = g$, a.e μ_L. Then f is S-integrable, by Theorem 3.12, Theorem 3.20(Remarks 3.21). If g is not bounded, let $B_n = g^{-1}([0, n])$, then , by the DCT;

$\int_X g d\mu_L = lim_{n\to\infty}(\int_X g\chi_{B_n} d\mu_L$

Using Theorem 3.4(ii), we can find a sequence $\{A_n\}_{n\in\mathcal{N}_{>0}}$ of disjoint sets, belonging to $\mathfrak{A}$, such that $\mu_L(B_n \setminus \bigcup_{i=1}^n A_i) < \frac{1}{n^2}.$, and with $A_n \subset B_n$. Now, considering $g\chi_{A_n}$, we can, by the above, find an S-integrable (bounded) $f_n \geq 0$, such that $^\circ f_n = g\chi_{A_n}$ a.e μ_L, and, without loss of generality, f_n is supported on A_n. Then;

$^\circ \int_X f_n d\nu = \int_X {^\circ f_n} d\mu_L = \int_{A_n} g d\mu_L$

$^\circ \int_X (\sum_{i=1}^n f_i) d\nu = \int_X (\sum_{i=1}^n {^\circ f_i}) d\nu = \int_{\bigcup_{i=1}^n A_i} g d\mu_L \leq \int_X f d\mu_L = C$

In particular;

$\int_X (\sum_{i=1}^n f_i) d\nu < C + \frac{1}{n}$

Moreover, as $g \leq n$ on B_n, and $\mu_L(B_n \setminus \bigcup_{i=1}^n A_i) < \frac{1}{n^2}$;

$\int_{B_n \setminus (\bigcup_{1 \leq i \leq n} A_i)} g d\mu_L < \frac{1}{n}$

$\int_{B_n} g d\mu_L - \int_X (\sum_{i=1}^n f_i) d\nu < \frac{1}{n}$ (†)

Now, extending the sequences $\{D_n\}_{n \in \mathcal{N}_{>0}}$ and $\{f'_n\}_{n \in \mathcal{N}_{>0}}$ in the usual way, where $f'_n = \sum_{i=1}^n f_i$, $D_n = \bigcup_{i=1}^n A_i$, we can find an infinite ω, such that $f'_\omega \geq 0$ is supported on D_ω, $f'_\omega | D_n = f'_n$, for $n \in \mathcal{N}_{>0}$, and;

$\int_X f'_\omega d\nu < C + \frac{1}{\omega}$ $(*)$

$\int_X f'_n d\nu \leq \int_X f'_\omega d\nu$

Letting $f = f_\omega$, we clearly have that ${}^\circ f = g$ a.e μ_L on $\bigcup_{i=1}^\infty A_i$, and $\mu_L(X - \bigcup_{i=1}^\infty A_i) = 0$. Therefore, ${}^\circ f = g$ a.e μ_L. By $(*)$, we obtain condition (i) of Definition 3.17. Now, given $\epsilon > 0$, choose n sufficiently large, such that;

$\int_X g \chi_{B_n} d\mu_L > C - \epsilon$ (††)

Then, by (†), (††), $(*)$;

$\int_X f'_n d\nu > C - \epsilon - \frac{1}{n}$

$\int_X (f - f'_n) d\nu < C + \frac{1}{\omega} - (C - \epsilon - \frac{1}{n}) = \epsilon + \frac{1}{n} + \frac{1}{\omega}$ (♯)

Now, it is easy to check condition (ii) of Definition 3.17, if $\nu(A) \simeq 0$, then;

$\int_A f d\nu = \int_A (f - f_n) d\nu + \int_A f'_n d\nu < \delta$

for any $\delta > 0$, by (♯) and condition (ii) of Definition 3.17 for f'_n. Hence, f is S-integrable, as required.

Case 2. $\mu_L(X) = \infty$. We do not actually require this result. In Section 5, when we consider spaces of infinite measure, one loses too much information if one replaces the transferred function by its lift. It should be clear to the reader by now, how the argument proceeds. As before, we can assume that g is supported on a σ-finite set $D = \bigcup_{n \in \mathcal{N}} B_n$, with $\mu_L(B_n) < \infty$. Now, using Theorem 3.4(i), and Case 1, a very similar argument produces the result. We leave the details to the reader.

□

The following definition can be found in [1];

Definition 3.32. *Let $\mathfrak{A}$ be a $^*\sigma$-algebra. For $1 \leq p < \infty$, let $SL^p(X, \mathfrak{A}, \nu)$ be the collection of equivalence classes of all $f : X \rightarrow^* \mathcal{R}$ such that f is $\mathfrak{A}$-measurable and $|f|^p$ is S-integrable, under the equivalence relation;*

$f_1 \sim f_2$ *iff* $(\int_X |f_1 - f_2|^p d\nu)^{\frac{1}{p}} \simeq 0$

We define a norm by $||f||_p = {}^\circ(\int_X |f|^p d\nu)^{\frac{1}{p}}$

Remarks 3.33. *One should check that this is a good definition, we leave this as an exercise. It relies on Minkowski's inequality.*

Following from this definition, we have the following theorem in [1];

Theorem 3.34. *Let $\mathfrak{A}$ be a $^*\sigma$-algebra and suppose $f : X \rightarrow^* \mathcal{R}$ is $\mathfrak{A}$-measurable. Then;*

(i). $f \in SL^p(X, \mathfrak{A}, \nu)$ iff ${}^\circ f \in L^p(X, \mathfrak{M}_L, \mu_L)$ and $||f||_p = ||{}^\circ f||_p$

(ii). If $g : X \rightarrow \mathcal{R}$ is in $L^p(X, \mathfrak{M}_L, \mu_L)$, there is a unique $f \in SL^p(X, \mathfrak{A}, \nu)$, such that ${}^\circ f = g$ a.e $d\mu_L$.

(iii). $SL^p(X, \mathfrak{A}, \nu)$ and $L^p(X, \mathfrak{M}_L, \mu_L)$ are isometrically isomorphic via the standard part mapping $f \rightarrow^\circ f$.

(iv). Suppose $\nu(X)$ is finite, $f \in SL^p(X, \mathfrak{A}, \nu)$, then, if $1 \leq q \leq p$, $f \in SL^q(X, \mathfrak{A}, \nu)$.

Remarks 3.35. *The proof can be found in* [1]. *$(i), (ii), (iii)$ are fairly routine, (iv) relies on the transfer of Holder's inequality.*

We require the following result in Section 6;

Theorem 3.36. *Suppose $g : X \rightarrow \mathcal{R}$ is integrable with respect to μ_L, $\mu_L(X) < \infty$, and $\epsilon > 0$ is standard, then there exist $F, G : X \rightarrow^* \mathcal{R}$, which are $\mathfrak{A}$-measurable, such that;*

(i). $G \leq g \leq F$.

(ii). $|\int_A g d\mu_L - \int_A G d\nu| < \epsilon$, $|\int_A g d\mu_L - \int_A F d\nu| < \epsilon$

for all $A \in \mathfrak{A}$.

Proof. Consider, first, the case when $g \geq 0$.

Upper Bound. As g is integrable, by Theorem 3.31, it has an S-integrable lifting F', such that ${}^\circ F' = g$ a.e μ_L, and;

$${}^\circ \int_X F' d\nu = \int_X g d\mu_L$$

Without loss of generality, we cam assume that $F' \geq 0$. Now let $\epsilon > 0$ be given and choose $\delta > 0$ such that $\mu_L(X)\delta < \frac{\epsilon}{2}$. Then $F' + \delta$ is S-integrable and $F' + \delta \geq f$ a.e μ_L, $(*)$, $F' + \delta > 0$. Moreover;

$${}^\circ \int_X (F' + \delta) d\nu = \int_X g d\mu_L + \delta\mu_L(X) < C + \frac{\epsilon}{2}, \ (**)$$

where $C = \int_X g d\mu_L$. Let $N \in \mathfrak{M}_L$, with $\mu_L(N) = 0$, such that $(*)$ holds on N^c. Let $N_n = N \cap g^{-1}((n-1, n])$, for $n \in \mathcal{N}_{>0}$, $N_0 = N \cap g^{-1}(0)$. Then $N = \bigcup_{n \geq 0} N_n$, and $\mu_L(N_n) = 0$. By Lemma 3.15 (3.4(i)), we can choose $U_n \supset N_n$, with $U_n \in \mathfrak{A}$, such that $\mu_L(U_n) < \frac{\epsilon}{4(n+1)^3}$. Inductively, define $F_0 = F' + \delta$, and, having defined F_n, let $F_{n+1} = F_n$ on U_{n+1}^c, and $F_{n+1} = F_n + n + 1$ on U_{n+1}. Then $\{F_n\}$ is an increasing sequence of $\mathfrak{A}$-measurable functions. Moreover;

$$\int_X F_{n+1} d\nu$$

$$= \int_{U_{n+1}^c} F_n d\nu + \int_{U_{n+1}} (F_n + (n+1)) d\nu$$

$$\simeq \int_X F_n d\nu + (n+1)\mu_L(U_{n+1})$$

$$< \int_X F_n d\nu + \frac{\epsilon}{4(n+1)^2}$$

$$\int_X F_n d\nu < C + \frac{\epsilon}{2} + \sum_{m=1}^{n} \frac{\epsilon}{4m^2} < C + \epsilon \text{ (using } (**))$$

We clearly have that for all $x \in N_n$, $g(x) \leq F_n$. Now, by countable comprehension, we can find an internal sequence $\{F_n\}_{n \in {}^*\mathcal{N}}$ extending the sequence $\{F_n\}_{n \in \mathcal{N}}$. By overflow, there exists an infinite ω, such that $F_n \leq F_\omega$, for all $n \in \mathcal{N}$, $F_\omega > 0$, and;

$\int_X F_\omega d\nu < C + \epsilon$, (†)

Clearly $g(x) \leq F_\omega(x)$, for all $x \in X$. Now, if $A \in \mathfrak{A}$, with;

$\int_A F_\omega d\nu - \int_A g d\mu_L > \epsilon$

then, using Theorem 3.16;

$\int_X F_\omega d\nu$

$= \int_A F_\omega d\nu + \int_{A^c} F_\omega d\nu$

$> \epsilon + \int_A g d\mu_L + \int_{A^c} g d\mu_L = C + \epsilon$

contradicting (†). Setting $F = F_\omega$ gives an upper bound.

Lower Bound. Again choose $\delta > 0$, with $\mu_L(X)\delta < \frac{\epsilon}{2}$. Let F' be as before, then $F' - \delta$ is S-integrable, $F' - \delta \leq g$ a.e μ_L, and:

$\int_X (F' - \delta) d\nu > C - \frac{\epsilon}{2}$

Again choose N, with $\mu_L(N) = 0$, such that $F' - \delta \leq g$ on N^c. Using Lemma 3.15(3.4(i)) again, we can choose a decreasing sequence of sets $\{U_n\}_{n \in \mathcal{N}_{>0}}$, belonging to $\mathfrak{A}$, with $U_n \supset N$, and $\mu_L(U_n) < \frac{1}{n}$. By S-integrability;

${}^\circ \int_{U_n} (F' - \delta) d\nu = \int_{U_n} {}^\circ(F' - \delta) d\mu_L$

and;

$lim_{n \to \infty} (\int_{U_n} {}^\circ(F' - \delta) d\mu_L) = 0$

by the DCT, as ${}^\circ(F' - \delta)\chi_{U_n}$ converges to 0 a.e μ_L. Hence, for sufficiently large n, we can assume that;

$\int_{U_n} (F' - \delta) d\nu < \frac{\epsilon}{2}$

Now let $G = (F' - \delta)$ on U_n^c, and $G = 0$ on U_n. Clearly $G(x) \leq g(x)$, for all $x \in X$. Moreover;

$\int_X Gd\nu$

$= \int_{U_n^c}(F' - \delta)d\nu$

$= \int_X(F' - \delta)d\nu - \int_{U_n}(F' - \delta)d\nu > C - \epsilon$

The same argument as above shows that, for all $A \in \mathfrak{A}$;

$\int_A gd\mu_L - \int_A Gd\nu \leq \epsilon$

Hence, G is a lower bound.

Now, if g is integrable μ_L, we can write $g = g^+ - g^-$, with $\{g^+, g^-\}$ integrable μ_L. Choosing $G \geq g^+$ and $H \leq g^-$, $G - H \geq (g^+ - g^-) = g$, choosing $G' \leq g^+$ and $H' \geq g^-$, $G' - H' \leq (g^+ - g^-) = g$, and, clearly, we can obtain the integral condition, using $\frac{\epsilon}{2}$.

□

Finally, we need the following technical lemmas, in Section 7;

Lemma 3.37. *Suppose $\nu(X) < \infty$, $f : X \rightarrow {}^*\mathcal{R}$ is $\mathfrak{A}$-measurable, and $g : X \rightarrow \mathcal{R}$ is $\mathfrak{M}_L$-measurable. Then, if;*

$^\circ\int_A f dv = \int_A gd\mu_L$

for all $A \in \mathfrak{A}$. Then $^\circ f = g$ a.e $d\mu_L$.

Proof. This is straightforward. We can easily reduce to the case when $g \geq 0$ and $f \geq 0$. Then use Theorem 3.16, to obtain a contradiction if there exists a set $A \in \mathfrak{A}$, for which $\mu_L(A)$ and $^\circ f > g$ on A, and use Theorem 3.12, if $^\circ f < g$, as we can assume that both $^\circ f$ and g are bounded on A. □

Lemma 3.38. *Given $(X, \mathfrak{A}, \nu)$ and $(X', \mathfrak{A}', \nu')$, with $\nu(X)$ and $\nu(X')$ finite, let $(X \times X', L(\mathfrak{A} \times \mathfrak{A}'), L(\nu \times \nu'))$ be the Loeb space corresponding to $(X \times X', \mathfrak{A} \times \mathfrak{A}', \nu \times \nu')$ and let $(X \times X', L(\mathfrak{A}) \times L(\mathfrak{A}'), L(\nu) \times L(\nu'))$ be the complete product, see* [37]*, of the Loeb spaces $(X, L(\mathfrak{A}), L(\nu))$ and $(X', L(\mathfrak{A}'), L(\nu'))$. Then;*

$L(\mathfrak{A} \times \mathfrak{A}') \supset L(\mathfrak{A}) \times L(\mathfrak{A}')$ *and* $L(\nu \times \nu')|L(\mathfrak{A}) \times L(\mathfrak{A}') = L(\nu) \times L(\nu')$

Proof. The proof can be found in [1]. □

The following result, see [2], is required in Section 4.

Theorem 3.39. *Suppose $\nu(X)$ is finite, and Y is a Hausdorff space with a countable base of open sets. If $g : X \rightarrow Y$ is $\mathfrak{M}_L$-measurable, there exists an internal $f : X \rightarrow {}^*Y$, which is $\mathfrak{A}$-measurable, such that ${}^\circ f(x) = g(x)$, a.e $\mu_L(x)$.*

Proof. The proof is in [2], Theorem 5.3.

□

Remarks 3.40. *A particularly simple application of this result is when $Y = \mathcal{C}$. However, all of the definitions and results given in this section extend easily to this case, by taking real and imaginary parts.*

4. Peano's Theorem

In this very short section, we give a proof of Peano's existence theorem, using the methods of nonstandard analysis. The theorem is the foundational result in dynamical systems. We let $\mathfrak{B}$ denote the completion of the Borel field on $[0,1] \times \mathcal{R}$ and $\mathcal{R}$, relative to μ, where μ denotes Lebesgue measure. See Definition 7.16, and Sections 5,6. We denote by $C[-c,c]$ the space of continuous functions on the closed interval $[-c,c]$, where $c \in \mathcal{R}$. We assume $C[-c,c]$ is equipped with the uniform norm, defined by $||g|| = max_{z\in[-c,c]}|g(x)|$. We let $\mathfrak{T}$ denote the topology on $C[-c,c]$ defined by $||\ ||$, and let $\mathfrak{D}$ denote the Borel field on $C[-c,c]$ generated by $\mathfrak{T}$. $U_{\epsilon,h} = \{H \in C[-c,c] : ||h - H|| < \epsilon\}$. $(Y, \mathfrak{C}, \lambda)$ will denote the hyperfinite interval, as given in Definition 7.16.

Theorem 4.1. *Peano's Theorem*

Let $g : [0,1] \times \mathcal{R} \rightarrow \mathcal{R}$ be bounded, measurable and continuous in the second variable. Let $x_0 \in \mathcal{R}$, then there exists a continuous function $x : [0,1] \rightarrow \mathcal{R}$, with $x(0) = x_0$, such that;

$x(t) = \int_0^t g(s, x(s))ds$ *for* $t \in [0,1]$

In particular $x(t)$ satisfies the differential equation;

$\frac{dx}{dt} = g(t, x(t))$, *with initial condition* $x(0) = x_0$, *a.e* μ

Remarks 4.2. *If g is continuous, then $\frac{dx}{dt} = g(t, x(t))$ everywhere on $[0,1]$ by the Fundamental Theorem of Calculus. If $x(t)$ satisfies the* *integral equation, then the second statement follows from the fact that g is absolutely continuous and Theorem 7.20 of* [37]*. If g is continuous on an open neighborhood of x_0 in $[0,1]\times\mathcal{R}$, then we automatically obtain the boundedness, measurability and second variable continuity conditions locally, and it is simple to adapt the following proof to obtain a local result. We leave this as an exercise for the reader.*

Proof. Without loss of generality $x_0 = 0$. Let $|g| \leq c$. Define $g_1 : [0,1] \rightarrow C([-c,c])$ by $g_1(t)(z) = g(t,z)$, for $|z| \leq c$. Then $g_1 : ([0,1],\mathfrak{B}) \rightarrow (C([-c,c]),\mathfrak{D})$ is measurable, $(*)$. To show $(*)$, as $\mathfrak{B}$ is a σ-algebra, it is sufficient to check that $g_1^{-1}(U_{\epsilon,h}) \in \mathfrak{B}$, for $U_{\epsilon,h} \in \mathfrak{T}$. As g is measurable, a.e $\mu(z)$ in $[-c,c]$, the functions $g_z(t)$ are measurable, see [37], Theorem 8.12. Now choose a countable dense subset W of $[-c,c]$ for which $g_w(t)$ is measurable for $w \in W$. Then, by continuity of g in the second variable, if $g_{t_0} = h$;

$$g_1^{-1}(U_{\epsilon,h}) = \bigcup_{n\in\mathcal{N}}\bigcap_{w\in W}\{t \in [0,1] : |g(t,w) - g(t_0,w)| \leq \epsilon - \tfrac{1}{n}\}$$

which belongs to $\mathfrak{B}$, hence $(*)$ is shown. Now define $g_2 : Y \rightarrow C[-c,c]$, by $g_2(t) = g_1({}^{\circ}t)$. Then, by Theorem 7.17, (the proof is in Section 5), $g_2 : (Y,\mathfrak{C}) \rightarrow (C[-c,c],\mathfrak{D})$ is measurable. Now, $C[-c,c]$ is separable in the uniform topology. This follows from the Stone-Weierstrass Theorem, see [9]. Moreover, as the topology is generated by a metric, the topology $\mathfrak{T}$ has a countable base of open sets. Clearly $C[-c,c]$ is Hausdorff. Therefore, by Theorem 3.39, there exists a measurable lifting $f_2 : (Y,\mathfrak{C}) \rightarrow ({}^*C[-c,c],{}^*\mathfrak{D})$ such that a.e $L(\lambda)(\tau)$ $f_2(\tau) \simeq g_2(\tau)$, $(**)$. Moreover, we can assume that ${}^*||f_2(\tau)|| \leq c$, for all $\tau \in Y$. This follows as the function $\Gamma : ({}^*C[-c,c],{}^*\mathfrak{D}) \rightarrow ({}^*C[-c,c],{}^*\mathfrak{D})$, defined by;

$$\Gamma(H) = \tfrac{c}{{}^*||H||}H \text{ if } {}^*||H|| > c$$

$$\Gamma(H) = H \text{ otherwise}$$

is measurable, by transfer. Then replacing f_2 by $\Gamma \circ f_2$, we still have measurabilty, and the condition $(**)$ still holds as ${}^*||g_2(\tau)|| \leq c$, for all $\tau \in Y$. Now define $f : Y \times {}^*[-c,c] \rightarrow {}^*\mathcal{R}$ by $f(\tau,X) = f_2(\tau)(X)$. Suppose ${}^*||f_\tau - g_{2,\tau}|| \simeq 0$ and $|Z| \leq c$, $(\dagger)$. $Z \simeq {}^{\circ}Z$ implies that $f_\tau(Z) \simeq f_\tau({}^{\circ}Z)$, this is a simple extension of Theorem 2.25. By the assumption $(\dagger)$, $f_\tau({}^{\circ}Z) \simeq g_{2,\tau}({}^{\circ}Z)$. By definition of g_1 and g_2, $g_{2,\tau}({}^{\circ}Z) = g_{1,{}^{\circ}\tau}({}^{\circ}Z) = g({}^{\circ}\tau,{}^{\circ}Z)$. So we have, a.e $L(\lambda)(\tau)$, for all $|Z| \leq c$, that $f(\tau,Z) \simeq g({}^{\circ}\tau,{}^{\circ}Z)$, $(\dagger\dagger)$, and, clearly, $|f| \leq c$. By Lemma 2.10, f is internal. We now define $X : {}^*[0,1] \rightarrow {}^*\mathcal{R}$, by first defining X_1 on the set $V=\{\frac{i}{\eta}:0\leq i<\eta\}$;

$X_1(0)=0$

$X_1(\frac{i+1}{\eta})=X_1(\frac{i}{\eta})+f(\frac{i}{\eta},X_1(\frac{i}{\eta}))\frac{1}{\eta}$, for $i<\eta-1$, $(***)$, ([5]).

and letting $X(\tau)=X_1(\frac{[\eta\tau]}{\eta})$, for $\tau\in Y$, $X(1)=X(\frac{\eta-1}{\eta})$. Again, by the transfer principle, we have that $X(\tau)\leq c\tau$ and $|X(\tau-\tau')|\leq c|\tau-\tau'|$, for $\tau,\tau'\in Y$. In particular, X is S-continuous in the sense of Definition 2.35 It follows, by Theorem 2.36, that we can define a continuous function $x:[0,1]\rightarrow\mathcal{R}$ by setting $x(t)={}^{\circ}X(\tau)$, for any $\tau\in{}^{*}[0,1]$, with $\tau\simeq t$. Then , by (††) and the definition of x, a.e $L(\lambda)$, $F_1(\tau)=f(\tau,X(\tau))\simeq g({}^{\circ}\tau,{}^{\circ}X(\tau))=g({}^{\circ}\tau,x({}^{\circ}\tau))$, (†††). Now F_1 is internal, so F, defined on Y by $F(\tau)=F_1(\frac{[\eta\tau]}{\eta})$ is $\mathfrak{C}$-measurable, and the condition (†††), still holds, using Theorem 3.4(ii). As F is bounded, it is S-integrable, by Theorem 3.30(i). Hence, it is a lift of the function $g_3:Y\rightarrow\mathcal{R}$, defined by $g_3(\tau)=g({}^{\circ}\tau,x({}^{\circ}\tau))$. Then, if $t\in[0,1]$;

$x(t)={}^{\circ}X(t)={}^{\circ}X_1(\frac{[\eta t]}{t})$

$={}^{\circ}(\frac{1}{\eta}{}^{*}\sum_{i=0}^{[\eta t]-1}f(\frac{i}{\eta},X_1(\frac{i}{\eta})))$

$=\int_0^{\frac{[\eta t]}{\eta}}Fd\lambda$

$=\int_0^{\frac{[\eta t]}{\eta}}g_3(\tau)dL(\lambda)(\tau)$

$=\int_0^t g(s,x(s))d\mu(s)$

by the definition of X_1 in $(***)$, the internal integral on $(Y,\mathfrak{C},\lambda)$, (see Remarks 3.10 and footnote 40), Theorem 3.20 and Theorem 7.17. Hence, the theorem is shown, ([6]).

[5]The existence and uniqueness of such a function follows immediately from the transfer principle. Clearly the statement;

$\forall n\in\mathcal{N}_{\geq 2}\forall f\in Q\exists!X_1\in R\forall m\in\mathcal{N}_{[0,n-2]}(X_1=0\wedge X_1(\frac{m+1}{n})=X_1(\frac{m}{n})+f(\frac{m}{n},X_1(\frac{m}{n}))\frac{1}{n})$

where $Q=\{f:[0,1)\times[-c,c]\rightarrow\mathcal{R}\}$, $R=\{X:V'\rightarrow R\}$, $V'=\{\frac{m}{n}:0\leq m\leq n-1\}$, holds in $\mathcal{R}$, hence, it holds in ${}^{*}\mathcal{R}$.

[6]One might try to adapt the method to prove the existence of solutions to more complicated ODE's. In fact, this was done to solve stochastic differential equations,

□

5. Fourier Analysis

In this section, we give a proof of the Fourier Inversion Theorem, using the methods of non-standard analysis. I became interested in this problem after reading through the notes [4]. We first require the following, which can be found in [40];

Definition 5.1. *We denote by the Schwartz space $S(\mathcal{R})$, the set of all functions $g : \mathcal{R} \rightarrow \mathcal{C}$, such that g and all its derivatives $\{g', g'', \ldots, g^{(n)}, \ldots, \}_{n\in\mathcal{N}}$ are rapidly decreasing, in the sense that;*

$sup_{x\in\mathcal{R}}|x|^k|g^{(n)}(x)| < \infty$. *(for all $k, n \geq 0$)*

For such a function g, we define its Fourier transform by;

$\hat{g}(t) = \int_{-\infty}^{\infty} g(x)e^{-\pi ixt}dx$

Remarks 5.2. *It is a well known fact that if $g \in S(\mathcal{R})$, then its Fourier transform $\hat{g} \in S(\mathcal{R})$ as well, see [40]. However, this is, perhaps, not the usual definition of the Fourier transform. In [40], it is given as;*

$\hat{g}(t) = \int_{-\infty}^{\infty} g(x)e^{-2\pi ixt}dx$

while, in [23], it is defined as;

$\hat{g}(t) = \int_{-\infty}^{\infty} g(x)e^{-ixt}dx$

Of course, these definitions only differ by a scaling factor, but for each one you choose, you get a distinct rescaled statement of the Inversion Theorem. Once you have proved the Fourier Inversion theorem for one definition, you obtain the other statements by a simple change of variables. The reason for our choice of notation will become apparent later.

Theorem 5.3. *Fourier Inversion Theorem*

---

see [22]. It is also an interesting question as to whether one can obtain solutions to PDE's this way. This was suggested to me by Mark Holland, who used a difference method in finding approximate solutions to the Black Scholes equation, see [10].

Let $g \in S(\mathcal{R})$, then;

$g(x) = \frac{1}{2}\int_{-\infty}^{\infty} \hat{g}(t)e^{\pi ixt}dt$ *for all* $x \in \mathcal{R}$.

Remarks 5.4. *There are many standard proofs of this result, for example in [40]. However, this is not the best statement possible. In [23], the requirement that $g \in S(\mathcal{R})$ is weakened to $g \in L^1(\mathcal{R}) \cap C$ and $\hat{g} \in L^1(\mathcal{R}) \cap C$, where C denotes the space of complex valued continuous functions on $\mathcal{R}$. In our proof, we do not actually require that $g \in S(\mathcal{R})$, but we do need some assumptions about the differentiability of g, and also about its rate of decrease. We have chosen this assumption, mainly because the Schwartz space seems to be often used in the presentation of the Fourier Inversion Theorem.*

We now introduce the principal spaces which we are going to work with;

Definition 5.5. *Let $\eta \in {}^*\mathcal{N} \setminus \mathcal{N}$, and $\omega \in {}^*\mathcal{N}$, with $\omega \geq n\eta$, for* all *$n \in \mathcal{N}$. We define;*

$\overline{\mathcal{R}}_{\omega,\eta} = \{\tau \in {}^*\mathcal{R} : -\frac{\omega}{\eta} \leq \tau < \frac{\omega}{\eta}\}$

We let $\mathfrak{C}$ be the $*$-finite algebra consisting of internal unions of intervals of the form $[\frac{i}{\eta}, \frac{i+1}{\eta})$, for $-\omega \leq i < \omega$.

We define a counting measure on $\mathfrak{C}$ by $\lambda([\frac{i}{\eta}, \frac{i+1}{\eta})) = \frac{1}{\eta}$.

Then $(\overline{\mathcal{R}}_{\omega,\eta}, \mathfrak{C}, \lambda)$ is a hyperfinite measure space with $\lambda(\overline{\mathcal{R}}_{\omega,\eta}) = \frac{2\omega}{\eta}$.

We denote by $(\overline{\mathcal{R}}_{\omega,\eta}, L(\mathfrak{C}), L(\lambda))$ the associated Loeb space, ([7]).

We let $(\mathcal{R}, \mathfrak{B}, \mu)$ denote the completion of the Borel field $\mathfrak{D}$ on $\mathcal{R}$, with respect to Lebesgue measure μ, ([8]).

[7] The existence of such a space follows from [24]. However, the uniqueness of the extension of $^\circ\lambda$ to $\sigma(\mathfrak{C})$ was only shown there in the case that λ is finite. Later, Ward Henson proved the uniqueness of the extended measure, even in the case that λ is infinite. After producing the extension, we are then passing to the completion, as in the remarks before Lemma 3.15.

[8] Again, Caratheodory's Theorem provides the existence of Lebesgue measure μ on the σ-algebra $\mathfrak{D}$ generated by the open sets. Uniqueness of the extension follows easily by restricting to finite intervals.

We let $\mathcal{R}^{+-\infty}$ denoted the extended real line $\mathcal{R}\cup\{+\infty,-\infty\}$, and let $\{g_\infty,\hat{g}_\infty\}$ be the extensions of functions in Definition 5.1, obtained by setting $g_\infty(+\infty)=g_\infty(-\infty)=0$, and similarly for $\hat{g}_\infty$.

Lemma 5.6. *There exists a unique σ-algebra $\mathfrak{B}'$ on $\mathcal{R}^{+-\infty}$, which separates the points $+\infty$ and $-\infty$, and such that $\mathfrak{B}'|_{\mathcal{R}}=\mathfrak{B}$. Moreover, there is a unique extension of μ to a measure μ' on $\mathfrak{B}'$ with the property that $\mu(\infty)=\mu(-\infty)=\infty$. The same holds with $\mathfrak{D}$ and $\mathfrak{D}'$ replacing $\mathfrak{B}$ and $\mathfrak{B}'$. The resulting measure space $(\mathcal{R}^{+-\infty},\mathfrak{B}',\mu')$ is the completion of $(\mathcal{R}^{+-\infty},\mathfrak{D}',\mu')$.*

Proof. The construction of $\mathfrak{B}'$ is easy. We let $\mathfrak{B}_{+\infty}$ consist of all sets of the form $B\cup\{+\infty\}$, where $B\in\mathfrak{B}$, and, similarly, define $\mathfrak{B}_{-\infty}$ and $\mathfrak{B}_{+-\infty}$. Then, let $\mathfrak{B}'=\mathfrak{B}\cup\mathfrak{B}_{+\infty}\cup\mathfrak{B}_{-\infty}\cup\mathfrak{B}_{+-\infty}$. Clearly, $\mathfrak{B}'$ separates the points $+\infty$ and $-\infty$, moreover $\mathfrak{B}'|_{\mathcal{R}}=\mathfrak{B}$. It is a simple exercise to verify that $\mathfrak{B}$ is a σ-algebra. In order to see uniqueness, let $\mathfrak{B}''$ have these properties. As $\mathfrak{B}''|_{\mathcal{R}}=\mathfrak{B}$, we have $\mathfrak{B}\subset\mathfrak{B}''$. Choose a set B containing $+\infty$, but not $-\infty$, then $\{+\infty\}=B\cap\bigcap_{n\in\mathcal{N}}(-n,n)^c$ belongs to $\mathfrak{B}''$. Moreover $\{+\infty,-\infty\}=\mathcal{R}^{+-\infty}\setminus\mathcal{R}$ belongs to $\mathfrak{B}''$, so, $-\infty$ belongs to $\mathfrak{B}''$. Hence, $\mathfrak{B}'\subset\mathfrak{B}''$. If C belongs to $\mathfrak{B}''$, then clearly $C\cap\mathcal{R}\in\mathfrak{B}$, so it must be of the above form, that is $\mathfrak{B}'=\mathfrak{B}''$. Now define μ' by setting $\mu=\mu'$ on $\mathfrak{B}$, and letting $\mu'(C)=\infty$, for any $C\in\mathfrak{B}'\setminus\mathfrak{B}$. It is straightforward to see that μ' defines a measure, with $\mu'(\infty)=\mu'(-\infty)=\infty$, extending μ. If μ'' satisfies these properties, then as any set $C\in\mathfrak{B}'\setminus\mathfrak{B}$ contains at least one of $\{+\infty,-\infty\}$, it must be ∞ on these sets, so $\mu'=\mu''$. Exactly the same argument gives the result for $\mathfrak{D}$ and $\mathfrak{D}'$. The completeness statement follows directly as $(\mathcal{R},\mathfrak{B},\mu)$ is complete, and any set of measure 0, μ', in $\mathfrak{B}'$, belongs to $\mathfrak{B}$.

□

Theorem 5.7. *The standard part mapping;*

$$st:(\overline{\mathcal{R}}_{\omega,\eta},L(\mathfrak{C}),L(\lambda))\rightarrow(\mathcal{R}^{+-\infty},\mathfrak{B}',\mu')$$

is measurable and measure preserving. In particular, if $\{g_\infty,\hat{g}_\infty\}$ are as in Definition 5.5, and $\{st^*(g_\infty),st^*(\hat{g}_\infty)\}$ are their pullbacks under st, then, $\{st^*(g_\infty),st^*(\hat{g}_\infty)\}$ are integrable with respect to $L(\lambda)$, $\{g_\infty,\hat{g}_\infty\}$ are integrable with respect to μ', $\{g,\hat{g}\}$ are integrable with respect to μ, and;

$$\int_{\overline{\mathcal{R}}_{\omega,\eta}}st^*(g_\infty)dL(\lambda)=\int_{\mathcal{R}^{+-\infty}}g_\infty d\mu'=\int_{\mathcal{R}}gd\mu$$

$\int_{\overline{\mathcal{R}}_{\omega,\eta}} st^*(\hat{g}_\infty)dL(\lambda) = \int_{\mathcal{R}+-\infty} \hat{g}_\infty d\mu' = \int_{\mathcal{R}} \hat{g} d\mu$

Proof. We let $\Sigma'_0 \subset \mathfrak{B}'$ denote the sets consisting of finite unions of the form;

$$[-\infty, b_1) \cup [a_2, b_2) \cup \ldots \cup [a_r, b_r) \cup [b_{r+1}, \infty]$$

where $b_1 \leq a_2 \ldots \leq b_{r+1}$ belong to $\mathcal{R}$. It is an easy exercise to check that Σ_0 is an algebra. Let $\mathfrak{D}' \subset \mathfrak{B}'$ be the σ-algebra generated by Σ'_0. Then $\mathfrak{D}'|_{\mathcal{R}}$ is just the Borel field $\mathfrak{D}$ on $\mathcal{R}$, and by Lemma 5.6, $\mathfrak{D}'$ is obtained from $\mathfrak{D}$ by adjoining at least one of the points $\{+\infty, -\infty\}$. Then $\mathfrak{B}'$ is just the completion of $\mathfrak{D}'$ with respect to $\mu'|_{\mathfrak{D}'}$, using the definition of $\mathfrak{B}$ and the fact that $\mathfrak{B}'|_{\mathcal{R}} = \mathfrak{B}$. Now, if $a, b \in \mathcal{R}$;

$$st^{-1}([a,b)) = \bigcup_{m=1}^{\infty} \bigcap_{n=1}^{\infty} [\tfrac{[\eta(a-\frac{1}{n})]}{\eta}, \tfrac{[\eta(b-\frac{1}{m})]}{\eta})$$

$$st^{-1}([-\infty,a)) = \bigcup_{m=1}^{\infty} [\tfrac{-\omega}{\eta}, \tfrac{[\eta(a-\frac{1}{m})]}{\eta}) \ (*)$$

where $[\ ]$ denotes integer part. Observing that $\{i \in {}^*\mathcal{Z} : -\omega \leq i \leq [\eta(b - \frac{1}{m})] - 1\}$ is internal, these sets belong to $L(\mathfrak{C})$. Now consider $\{B \in \mathfrak{B}' : st^{-1}(B) \in L(\mathfrak{C})\}$. This is a σ-algebra containing $\mathfrak{D}'$ by $(*)$. In particular, $st^{-1}(-\infty)$ and $st^{-1}(+\infty)$ belong to $L(\mathfrak{C})$. Moreover;

$$L(\lambda)(st^{-1}([a,b))) = lim_{m\rightarrow\infty} lim_{n\rightarrow\infty} {}^\circ(b - a + \tfrac{1}{n} - \tfrac{1}{m}) = (b-a)$$

$$L(\lambda)(st^{-1}(+\infty)) = L(\lambda)(st^{-1}(-\infty)) = \infty \ (**)$$

In the first claim, we have used elementary properties of measures on σ-algebras and the definition of $\lambda|\mathfrak{C}$. In the second claim, we have used the fact that $st^{-1}(+\infty) \supset [\frac{\omega}{2\eta}, \frac{\omega}{\eta})$, and $L(\lambda)([\frac{\omega}{2\eta}, \frac{\omega}{\eta})) = {}^\circ(\frac{\omega}{2\eta}) = \infty$, by the choice of ω. Similarly, for $st^{-1}(-\infty)$. It follows that the push forward measure $st_*(L(\lambda))$ on $\mathfrak{D}'$, agrees with μ on the algebra $\Sigma_0|\mathcal{R}$, hence, by footnote 8, it agrees with μ on $\mathfrak{D} = \mathfrak{D}'|_{\mathcal{R}}$. By Lemma 5.6, it agrees with μ' on $\mathfrak{D}'$. Now if $B \in \mathfrak{B}'$, we can find $C \subset B \subset D$, with C and D belonging to $\mathfrak{D}'$, such that $\mu'(D \setminus C) = 0$. Then $st^{-1}(C) \subset st^{-1}(B) \subset st^{-1}(D)$ and $L(\lambda)(st^{-1}(D \setminus C) = L(\lambda)(st^{-1}(D) \setminus st^{-1}(C)) = 0$. Hence, as $(\overline{\mathcal{R}}_{\omega,\eta}, L(\mathfrak{C}), L(\lambda))$ is complete, we have that $st^{-1}(B) \in L(\mathfrak{C})$ and $L(\lambda)(st^{-1}(B)) = L(\lambda)(st^{-1}C) = \mu'(C) = \mu'(B)$, as required. For the second part of the theorem, observe that $S(\mathcal{R}) \subset L^1(\mathcal{R})$ and use Remarks 5.2. Clearly, the extensions $\{g_\infty, \hat{g}_\infty\}$ are $\mathfrak{D}'$-measurable, and, using [37], Definition 1.23, and the proof of Lemma 3.15;

$\int_{\mathcal{R}+-\infty} g_\infty d\mu' = \int_{\mathcal{R}} g_\infty d\mu' + \int_{+\infty,-\infty} g_\infty d\mu' = \int_{\mathcal{R}} g d\mu$

and, similarly, for $\hat{g}$. Then, it follows, using the first part of the theorem, and the change of variables formula, see footnote 15, that, $\{st^*(g_\infty), st^*(\hat{g}_\infty)\}$ are integrable with respect to $L(\lambda)$, and;

$\int_{\overline{\mathcal{R}}_{\omega,\eta}} st^*(g_\infty) dL(\lambda) = \int_{\mathcal{R}+-\infty} g_\infty d\mu'$

and, similarly, for $st^*(\hat{g}_\infty)$.

□

We make the following;

Definition 5.8. *Let $(G,+,0)$ be a finite commutative group, and let $(\mathcal{C}^*,\centerdot,1)$ denote the multiplicative group of complex numbers, with absolute value 1, then by a a character γ of G, we mean a homomorphism $\gamma : G \rightarrow \mathcal{C}^*$.*

Let $m, n \in \mathcal{N}_{>0}$. We let $(\mathcal{Z}_m,+,0) = (\mathcal{Z}/m\mathcal{Z},+,0)$ denote the additive group of integers mod m. For $x, y \in \mathcal{Z}_m$, we let xy denote ordinary multiplication in $\mathcal{Z}$, where $\{x, y\}$ are uniquely represented in $\{0,\ldots,m-1\}$

$G_{2m} = \{-m, -(m-1), \ldots, m-1\}$ denotes the group of order $2m$, with addition given by $m_1 + m_2 = S^{m_1}(m_2)$, where S is the shift map $S(x) = x+1$ if $x \neq m-1$, $S(m-1) = -m$.

$G_{m,n} = \{\frac{-m}{n}, \frac{-(m-1)}{n}, \ldots, \frac{m-1}{n}\}$ denotes the group of order $2m$, with addition as defined for G_m. As before, for $x \in G_{m,n}$,$y \in G_{m,n}$ or $y \in \mathcal{Z}$, we let xy denote ordinary multiplication in $\mathcal{Z}$.

For a finite commutative group G, we let $\mathfrak{G}$ denote the finite σ-algebra consisting of all subsets of G, and μ_G the associated probability measure. $L^1(G)$ denotes the set of functions $g : G \rightarrow \mathcal{C}$. For $g, h \in L^1(G)$, we let $< g, h >= \int_G g\bar{h} d\mu_G$.

The following can be found in [25];

Theorem 5.9. *Let $(G,+,0)$ be a finite commutative group of order m, then there exist exactly m characters on G, and they form an orthonormal basis of $L^1(G)$, with respect to $<,>$, ([9]). The characters on $\mathcal{Z}_m$ are given by;*

$\gamma_k(x) = exp(\frac{2\pi i}{m}kx)$ *for* $k \in \{0, 1, \ldots, m-1\}$

Definition 5.10. *Let $(G,+,0)$ be a finite commutative group of order m, and let G_* denote its commutative group of characters, of order m, ([10]), then, if $g \in L^1(G)$, we define $\hat{g} : G_* \rightarrow \mathcal{C}$, by;*

$\hat{g}(\gamma) =< g, \gamma >= \int_G g\bar{\gamma} d\mu_G$

We then obtain;

Theorem 5.11. *Inversion Theorem for Finite Groups*

Let $\{G, G_*, g, \hat{g}\}$ be as in Definition 5.10, then;

$g(x) = \sum_{j=0}^{m-1} \hat{g}(\gamma_j)\gamma_j(x)$

where $x \in G$, and j enumerates G_*.

Proof. This is almost immediate. By Theorem 5.9;

$g = \sum_{j=0}^{m-1} < g, \gamma_j > \gamma_j$ in $L^1(G)$

Then, by Definition 5.10, and the fact that $\mu_G(x) > 0$, if $x \in G$;

$g(x) = \sum_{j=0}^{m-1} < g, \gamma_j > \gamma_j(x) = \sum_{j=0}^{m-1} \hat{g}(\gamma_j)$

□

We now compute the character group on $G_{m,n}$;

[9]It is shown in [25] that the characters form an orthogonal basis with respect to the measure $m\mu_G$. However, it is then a simple computation, using the definition of a character, to see that they are an orthonormal basis with respect to the probability measure μ_G

[10]In fact, G and G_* are isomorphic, see [25].

Lemma 5.12. *Let $G_{m,n}$ be as in Definition 5.8, then the characters on $G_{m,n}$ are given by;*

$\gamma_y(x) = exp(\frac{\pi i n^2}{m} xy)$

where $x, y \in G_{m,n}$.

Proof. First observe that there exists an isomorphism $\phi : G_m \rightarrow \mathcal{Z}_{2m}$, defined by $\phi(x) = (x + 2m)_{mod2m}$. Hence, by Theorem 5.9, the characters on G_m are given by;

$exp(\frac{2\pi i}{2m}(x + 2m)_{mod2m} j) = exp(\frac{\pi i}{m}(x + 2m)_{mod2m} j) = exp(\frac{\pi i}{m} xj)$

where $x \in G_m$, $j \in \{0, 1, \ldots, 2m - 1\}$. Here, we have also used the facts that;

$\frac{[x+2m]_{mod2m}}{m} = \frac{x}{m}$, if $0 \leq x \leq m - 1$

$\frac{[x+2m]_{mod2m}}{m} = \frac{x}{m} + 2$, if $-m \leq x < 0$

and $exp(2\pi i) = 1$. Now writing $j = y + m$, for $y \in G_m$, we obtain that;

$exp(\frac{\pi i}{m} xj) = -exp(\frac{\pi i}{m} y) = exp(\frac{\pi i}{m}(y - m))$

Observe that the characters $exp(\frac{\pi i}{m}(y - m))$ correspond to $exp(\frac{\pi i}{m} y')$, where $y' = y - m$ belongs to $\{-m, \ldots, -1\}$ if $y \in \{0, \ldots, m - 1\}$, and correspond to $exp(\frac{\pi i}{m} y'')$, where $y'' = y + m$ belongs to $\{0, \ldots, m - 1\}$ if $y \in \{-m, \ldots, -1\}$. Hence, the characters in G_{m*} are given by;

$\gamma_y(x) = exp(\frac{\pi i}{m} xy)$ $(*)$

for $x, y \in G_m$. Now observe there exists an isomorphism $\psi : G_{m,n} \rightarrow G_m$ defined by $\psi(x) = nx$. Hence, by $(*)$, the characters in $G_{m,n*}$ are given by;

$\gamma_y(x) = exp(\frac{\pi i}{m}(nx)(ny)) = exp(\frac{\pi i n^2}{m} xy)$

for $x, y \in G_{m,n}$. □

Definition 5.13. *Let $n \in \mathcal{N}_{>0}$, let $G_{n^2,n}$ be the group of order $2n^2$, as in Definition 5.8, and let $g \in L^1(G_{n^2,n})$. Let $\mathfrak{G}$ be as before, and let λ_G be the rescaled measure, given by $\lambda_G = 2n\mu_G$. Then, we define*

$\hat{g} \in L^1(G_{n^2,n})$ *to be the function;*

$\hat{g}(t) = \int_{G_{n^2,n}} g(x)exp(-\pi ixt)d\lambda_G \; (t \in G_{n^2,n}, x \in G_{n^2,n})$

Theorem 5.14. *Inversion Theorem for* $G_{n^2,n}$

Let $\{G_{n^2,n}, \nu_G, g, \hat{g}\}$ *be as in Definition 5.13, then;*

$g(x) = \frac{1}{2}\int_{G_{n^2,n}} \hat{g}(t)exp(\pi ixt)d\lambda_G \; (x \in G_{n^2,n})$

Proof. By Lemma 5.12, the characters on $G_{n^2,n}$ are given by;

$\gamma_y(x) = exp(\frac{\pi in^2}{n^2}xy) = exp(\pi ixy) \; (*)$

for $x, y \in G_{n^2,n}$. Using Definition 5.10, and the fact that $\mu_G(x) = \frac{1}{2n^2}$, for $x \in G_{n^2,n}$, we have;

$\hat{g}(\gamma_y) = \frac{1}{2n^2}\sum_{k=-n^2}^{n^2-1} g(\frac{k}{n})exp(-\pi i\frac{k}{n}y) \; (**)$

where $y \in G_{n^2,n}$. By Theorem 5.11, $(*), (**)$ and the fact that $\lambda_G(x) = \frac{1}{n}$, for $x \in G_{n^2,n}$;

$g(x) = \sum_{l=-n^2}^{n^2-1} \hat{g}(\gamma_{\frac{l}{n}})\gamma_{\frac{l}{n}}(x)$

$= \sum_{l=-n^2}^{n^2-1} \hat{g}(\gamma_{\frac{l}{n}})exp(\pi i\frac{lx}{n})$

$= \sum_{l=-n^2}^{n^2-1}[\frac{1}{2n^2}\sum_{k=-n^2}^{n^2-1} g(\frac{k}{n})exp(-\pi i\frac{k}{n}\frac{l}{n})]exp(\pi i\frac{lx}{n})$

$= \frac{1}{2}\frac{1}{n}\sum_{l=-n^2}^{n^2-1}[\frac{1}{n}\sum_{k=-n^2}^{n^2-1} g(\frac{k}{n})exp(-\pi i\frac{k}{n}\frac{l}{n})]exp(\pi i\frac{lx}{n})$

$= \frac{1}{2}\frac{1}{n}\sum_{l=-n^2}^{n^2-1}[\int_{G_{n^2,n}} g(y)exp(-\pi iy\frac{l}{n})d\lambda_G]exp(\pi i\frac{lx}{n})$

$= \frac{1}{2}\frac{1}{n}\sum_{l=-n^2}^{n^2-1} \hat{g}(\frac{l}{n})exp(\pi i\frac{lx}{n})$

$= \frac{1}{2}\int_{G_{n^2,n}} \hat{g}(t)exp(\pi ixt)d\lambda_G$

□

Definition 5.15. *We let* $\overline{\mathcal{R}_\eta} = \overline{\mathcal{R}_{\eta^2,\eta}}$ *and let* $\{\mathfrak{C}_\eta, \lambda_\eta\}$ *be as before. We let* $\mathfrak{C}_\eta^2$ *denote the* **-finite algebra on* $\overline{\mathcal{R}_\eta}^2$*, consisting of internal unions of the form* $[\frac{k}{\eta}, \frac{k+1}{\eta}) \times [\frac{j}{\eta}, \frac{j+1}{\eta})$, $-\eta^2 \leq k, j < \eta^2$*, and* λ_η^2 *be the counting* *measure on $\mathfrak{C}_\eta^2$, defined by $\lambda_\eta^2([\frac{k}{\eta},\frac{k+1}{\eta})\times[\frac{j}{\eta},\frac{j+1}{\eta}))=\frac{1}{\eta^2}$.*

We let ${}^*exp(\pi ixt), {}^*exp(-\pi ixt) : {}^*\mathcal{R}^2 \rightarrow {}^*\mathcal{C}$ be the transfers of the functions $exp(\pi ixt), exp(-\pi ixt) : \mathcal{R}^2 \rightarrow \mathcal{C}$, and use the same notation to denote the restrictions of the transfers to $\overline{\mathcal{R}_\eta}^2$.

We let $exp_\eta(\pi ixt), exp_\eta(-\pi ixt) : \overline{\mathcal{R}_\eta}^2 \rightarrow {}^*\mathcal{C}$ denote their $\mathfrak{C}_\eta^2$-measurable counterparts, defined by;

$exp_\eta(\pi ixt) = {}^*exp(\pi i\frac{[\eta x]}{\eta}\frac{[\eta t]}{\eta})$, $(x,t)\in\overline{\mathcal{R}_\eta}^2$

and, similarly, for $exp_\eta(-\pi ixt)$. Given $f : \overline{\mathcal{R}_\eta} \rightarrow {}^*\mathcal{C}$, which is $\mathfrak{C}_\eta$-measurable, we define;

$\hat{f}_\eta(t) = \int_{\overline{\mathcal{R}_\eta}} f(x)exp_\eta(-\pi ixt)d\lambda_\eta$

so $\hat{f}_\eta : \overline{\mathcal{R}_\eta} \rightarrow {}^*\mathcal{C}$ is $\mathfrak{C}_\eta$-measurable. $(*)$

Given $g : \mathcal{R} \rightarrow \mathcal{C}$, we let ${}^*g : {}^*\mathcal{R} \rightarrow {}^*\mathcal{C}$ denote its transfer and its restriction to $\overline{\mathcal{R}_\eta}$. We let g_η denote its $\mathfrak{C}_\eta$-measurable counterpart, as above, and let $\hat{g}_\eta$ be as in $(*)$.

For $n\in\mathcal{N}$, we let $\mathcal{R}_n = \overline{\mathcal{R}_n}\cap\mathcal{R}$. We let $\mathfrak{C}_{n,st}$ consist of all finite unions of intervals of the form $[\frac{i}{n},\frac{i+1}{n})$, for $-n^2\leq i\leq n^2-1$. $\lambda_{n,st}$ is defined on $\mathfrak{C}_{n,st}$, by setting $\lambda_n([\frac{i}{n},\frac{i+1}{n}))=\frac{1}{n}$.

$\{\mathfrak{C}_{n,st}^2, \lambda_{n,st}^2, exp_{n,st}(\pi ixt), exp_{n,st}(-\pi ixt)\}$ are all defined as above, restricting to $\mathcal{R}$. If $g : \mathcal{R} \rightarrow \mathcal{C}$, we similarly define, $\{g_{n,st}, \hat{g}_{n,st}\}$, (st is suggestive notation for standard). Observe that $\lambda_{n,st}$ is just the restriction of Lebesgue measure μ to $\mathfrak{C}_{n,st}$, and transfers to λ_n.

$\{exp_{n,st}(\pi ixt), exp_{n,st}(-\pi ixt), g_{n,st}, \hat{g}_{n,st}\}$ are all standard functions, which transfer to $\{exp_n(\pi ixt), exp_n(-\pi ixt), g_n, \hat{g}_n\}$.

Finally, we let $\mathfrak{C}_{n,ext}$ denote the σ-algebra on $\mathcal{R}$, consisting of countable unions of intervals of the form $[\frac{i}{n},\frac{i+1}{n})$, for $i\in\mathcal{Z}$, and $\lambda_{n,ext}$ be the corresponding measure. We similarly define $\{\mathfrak{C}_{n,ext}^2, \lambda_{n,ext}^2, exp_{n,ext}(\pi ixt), exp_{n,ext}(-\pi ixt)\}$

If $g : \mathcal{R} \rightarrow \mathcal{C}$, we let $g_{n,ext} : \mathcal{R} \rightarrow \mathcal{C}$ be the $\mathfrak{C}_{n,ext}$-measurable function obtained by setting $g_{n,ext}(x) = g(\frac{[nx]}{n})$, so $g_{n,ext}|_{\mathcal{R}_n} = g_{n,st}$.

We now have;

Lemma 5.16. *Inversion Theorem for* $\overline{\mathcal{R}_\eta}$

Let $\{\overline{\mathcal{R}_\eta}, \lambda_\eta, f, \hat{f}_\eta\}$ *be as in Definition 5.15, then;*

$f(x) = \frac{1}{2}\int_{\overline{\mathcal{R}_\eta}} \hat{f}_\eta(t) exp_\eta(\pi ixt) d\lambda_\eta(t)\ (x \in \overline{\mathcal{R}_\eta})$

Proof. As $f(x)$ is $\mathfrak{C}_\eta$-measurable and $exp_\eta(\pi ixt)$ is $\mathfrak{C}_\eta^2$-measurable, both sides of the equation are unchanged if we replace x by $\frac{[\eta x]}{\eta}$. Now the result follows directly, by transfer, from the corresponding result for $G_{n^2,n}$, Theorem 5.14, and the definition of the internal integral $\int_{\overline{\mathcal{R}_\eta}}$ on $\overline{\mathcal{R}_\eta}$, see Remarks 3.10 and footnote 46,([11]). □

We now want to specialise the result of Lemma 5.16 to $(\overline{\mathcal{R}_\eta}, L(\mathfrak{C}_\eta), L(\lambda_\eta))$, using the results of Section 3. The problem now is to obtain the S-integrability conditions.

Theorem 5.17. *Let* $g \in S(\mathcal{R})$*, then* g_η*, as given in Definition 5.15, is* S*-integrable on* $\overline{\mathcal{R}_\eta}$*, in the sense of Definition 3.26. Moreover* ${}^\circ g_\eta = st^*(g_\infty)$*, everywhere* $L(\lambda_\eta)$*, and;*

${}^\circ\int_{\overline{\mathcal{R}_\eta}} g_\eta d\lambda_\eta = \int_{\overline{\mathcal{R}_\eta}} st^*(g_\infty) dL(\lambda_\eta) = \int_{\mathcal{R}} g d\mu$

Proof. We first claim that $g_{n,ext}$ is integrable μ, and $lim_{n\rightarrow\infty}||g-g_{n,ext}||_{L^1} = 0$, $(*)$. In order to see this, let $\epsilon > 0$ be standard, and choose $N \in \mathcal{N} \geq 2$, such that;

$\int_{-\infty}^{\infty} |g| d\mu - \int_{-N}^{N} |g| d\mu < \frac{\epsilon}{3}$

and $N > \frac{9C}{\epsilon}$. As $|g|$ is continuous on the interval $[-N, N]$, by Darboux's theorem, see [3], there exists $M \in \mathcal{N}$, such that for all $n \geq M$;

$\int_{-N}^{N} (|g - g_{n,ext}|) d\mu < \frac{\epsilon}{3}$

Now, for $n \in \mathcal{N}_{>0}$, using Definition 5.1;

[11]If the reader is anxious about some ambiguity in transferring double sums or integrals, the important point to realise is that the * operator factors through any set of standard predicates or functions, so ${}^*\mathcal{R} \models (\forall n \in {}^*\mathcal{N})({}^*(S_{1,n} \circ S_{2,n}) = ({}^*S_{1,n} \circ {}^*S_{2,n}))$ if $\{{}^*S_{1,n}, {}^*S_{2,n}\}$ are hyperfinite sums, see Definition 2.19.

$\int_{|x|>N}|g_{n,ext}|(x)d\mu(x)$

$=\frac{1}{n}(\sum_{|j|\geq Nn+1}|g(\frac{j}{n})|+|g(N)|)$

$\leq|\frac{g(N)}{n}|+\frac{1}{n}\sum_{|j|\geq Nn+1}\frac{Cn^2}{j^2}$

$\leq\frac{C}{N}+Cn\int_{|x|>Nn}\frac{1}{x^2}dx$

$=\frac{C}{N}+\frac{2Cn}{Nn}=\frac{3C}{N}<\frac{\epsilon}{3}$

Combining these estimates, it follows that, $g_{n,ext}$ is integrable μ, and for $n\geq M$;

$\int_{-\infty}^{\infty}|g-g_{n,ext}|d\mu<\epsilon$

As ϵ was arbitrary, we obtain the result $(*)$. Now, using $(*)$, choose $N_1\in\mathcal{N}$, such that $||g\chi_{[L,N]}||_{L^1}<\frac{\epsilon}{2}$ and $||g-g_{n,ext}||_{L^1}<\frac{\epsilon}{2}$, for all $n\in\mathcal{N}_{>0}$, and $L,N\in\mathcal{Z}$, $LN\geq 0$,with $min(n,|L|,|N|)>N_1$. Then;

$||g_{n,ext}\chi_{[L,N]}||_{L^1}\leq||(g_{n,ext}-g)\chi_{[L,N]}||_{L^2}+||g\chi_{L,N}||_{L^1}<\epsilon\ (**)$

for all such $\{n,L,N\}$. We now transfer the result $(**)$. We have that;

$\mathcal{R}\models(\forall n_{(n>N_1)})(\forall L,N_{(LN\geq 0,N_1<|L|,|N|<n)})\int_L^N|g_{n,st}|d\lambda_{n,st}<\epsilon$

Hence, the corresponding statement is true in ${}^*\mathcal{R}$. In particular, if η is infinite, and $\{L,N\}$ are infinite, of the same sign, belonging to $\overline{\mathcal{R}}_\eta$, we have that;

$\int_L^N|g_\eta|d\lambda_\eta<\epsilon$

As ϵ was arbitrary we conclude that;

$\int_L^N|g_\eta|d\lambda_\eta\simeq 0\ (***)$

for all infinite $\{L,N\}$, of the same sign, in $\overline{\mathcal{R}}_\eta$. Now consider the internal sequence;

$\{s_n\}_{1\leq n\leq\eta}=\{\int_{\overline{\mathcal{R}}_\eta}(|g_\eta-g_\eta\chi_{[-n,n)}|)d\lambda_\eta\}_{1\leq n\leq\eta}$

Then, by $(***)$, $s_{\omega'} \simeq 0$, for all infinite $\omega' \leq \eta$. Applying Theorem 2.23, we have that $lim_{n\rightarrow\infty}({}^\circ s_n) = 0$. That is;

$$lim_{n\rightarrow\infty}{}^\circ(\int_{\overline{\mathcal{R}_\eta}} |g_\eta - g_\eta\chi_{[-n,n)}|d\lambda_\eta) = 0 \ (\dagger)$$

As g is bounded by M, the same is true for g_η, hence, the functions $\{g_\eta\chi_{[-n,n)}\}$ are finite, in the sense of Definition 3.29. Applying Theorem 3.30(ii) and (†), we obtain that g_η is S-integrable. Now, using the fact that $lim_{x\rightarrow\infty}g(x) = 0$, it is straightforward, using Theorem 2.22, to show that $g_\eta(x) \simeq 0$, for all infinite $x \in \overline{\mathcal{R}_\eta}$. As g is continuous, by Theorem 2.25, we have that $g_\eta(x) = {}^*g(\frac{[\eta x]}{\eta}) \simeq g({}^\circ x)$, for all finite $x \in \overline{\mathcal{R}_\eta}$. Hence, for all $x \in \overline{\mathcal{R}_\eta}$, ${}^\circ g_\eta(x) = st^*(g_\infty)(x)$. Finally, by Theorem 3.24 and Theorem 5.7;

$${}^\circ \int_{\overline{\mathcal{R}_\eta}} g_\eta d\lambda_\eta = \int_{\overline{\mathcal{R}_\eta}} st^*(g_\infty)dL(\lambda_\eta) = \int_{\mathcal{R}} gd\mu$$

□

The corresponding result for $\hat{g}_\eta$ is more difficult to show. We require the following;

Definition 5.18. *If $n \in \mathcal{N}$, and $g_{n,st}$ is $\mathfrak{C}_{n,st}$-measurable, we define the discrete derivative $g'_{n,st}$ by;*

$$g'_{n,st}(\tfrac{j}{n}) = n(g_{n,st}(\tfrac{j+1}{n}) - g_{n,st}(\tfrac{j}{n})) \ (-n^2 \leq j < n^2 - 1)$$

$$g'_{n,st}(\tfrac{n^2-1}{n}) = 0$$

$$g'_{n,st}(x) = g'_{n,st}(\tfrac{[nx]}{n}) \ (x \in \mathcal{R}_n)$$

and the shift $g^{sh}_{n,st}$ by;

$$g^{sh}_{n,st}(\tfrac{j}{n}) = g_{n,st}(\tfrac{j+1}{n}) \ (-n^2 \leq j < n^2 - 1)$$

$$g^{sh}_{n,st}(\tfrac{n^2-1}{n}) = 0$$

$$g^{sh}_{n,st}(x) = g^{sh}_{n,st}(\tfrac{[nx]}{n}) \ (x \in \mathcal{R}_n)$$

So both are $\mathfrak{C}_{n,st}$-measurable.

Lemma 5.19. *Discrete Calculus Lemmas*

Let $\{g_{n,st}, h_{n,st}\}$ *be* $\mathfrak{C}_{n,st}$*-measurable and let* $\{g'_{n,st}, h'_{n,st}, g^{sh}_{n,st}, h^{sh}_{n,st}\}$ *be as in Definition 5.18. Then;*

(i). $\int_{\mathcal{R}_n} g'_{n,st} d\lambda_{n,st} = g_{n,st}(\frac{n^2-1}{n}) - g_{n,st}(-n)$

(ii). $(g_{n,st}h_{n,st})' = g'_{n,st}h^{sh}_{n,st} + g_{n,st}h'_{n,st}$

(iii). $\int_{\mathcal{R}_n} g'_{n,st}h_{n,st} d\lambda_{n,st} = -\int_{\mathcal{R}_n} g^{sh}_{n,st}h'_{n,st} d\lambda_{n,st} + gh_{n,st}(\frac{n^2-1}{n}) - gh_{n,st}(-n)$

Proof. (i). We have, using Definition 5.18, see also Remarks 3.10;

$\int_{\mathcal{R}_n} g'_{n,st} d\lambda_{n,st}$

$= \frac{1}{n}\sum_{j=-n^2}^{n^2-2} g'_{n,st}(\frac{j}{n})$

$= \frac{1}{n}\sum_{j=-n^2}^{n^2-2} n(g_{n,st}(\frac{j+1}{n}) - g_{n,st}(\frac{j}{n}))$

$= g_{n,st}(\frac{n^2-1}{n}) - g_{n,st}(-n)$

(ii). Again, by Definition 5.18;

$(gh_{n,st})'(\frac{j}{n})$

$= n(gh_{n,st}(\frac{j+1}{n}) - gh_{n,st}(\frac{j}{n}))$

$= n((g_{n,st}(\frac{j+1}{n}) - g_{n,st}(\frac{j}{n}))h_{n,st}(\frac{j+1}{n}) + g_{n,st}(\frac{j}{n})(h_{n,st}(\frac{j+1}{n}) - h_{n,st}(\frac{j}{n})))$

$= g'_{n,st}(\frac{j}{n})h^{sh}_{n,st}(\frac{j}{n}) + g_{n,st}(\frac{j}{n})h'_{n,st}(\frac{j}{n})$

$= (g'_{n,st}h^{sh}_{n,st} + g_{n,st}h'_{n,st})(\frac{j}{n})$ $(-n^2 \leq j \leq n^2-2)$

$(g'_{n,st}h^{sh}_{n,st} + g_{n,st}h'_{n,st})(\frac{n^2-1}{n}) = (g_{n,st}h_{n,st})'(\frac{n^2-1}{n}) = 0$

(iii). By $(i), (ii)$;

$\int_{\mathcal{R}_n} (h_{n,st}g_{n,st})' d\lambda_{n,st}$

$= \int_{\mathcal{R}_n} (h'_{n,st}g^{sh}_{n,st} + h_{n,st}g'_{n,st}) d\lambda_{n,st}$

$= gh_{n,st}(\frac{n^2-1}{n}) - gh_{n,st}(-n)$

□

Definition 5.20. *For $n \in \mathcal{N}$, we let $\theta_n : \mathcal{R} \rightarrow \mathcal{C}$ be defined by $\theta_n(t) = n(exp(\frac{-\pi it}{n}) - 1)$, and let $\beta_n : \mathcal{R} \rightarrow \mathcal{C}$ be defined by $\beta_n(t) = n(exp(\frac{\pi it}{n}) - 1)$. We let $\{\phi_n, \psi_n\}$ denote their $\mathfrak{C}_n$-measurable counterparts on $\mathcal{R}_n$. If $g_{n,st}$ is $\mathfrak{C}_{n,st}$-measurable, we let;*

$$C_n(t) = g_{n,st}(\tfrac{n^2-1}{n})exp_{n,st}(-\pi i\tfrac{n^2-1}{n}t) - g_{n,st}(-n)exp_{n,st}(-\pi i(-n)t)$$

$$D_n(t) = -\tfrac{1}{n}g_{n,st}(-n)exp_{n,st}(\pi i\tfrac{t}{n})exp_{n,st}(-\pi i(-n)t).$$

$$C'_n(t) = -g'_{n,st}(-n)exp_{n,st}(-\pi i(-n)t)$$

$$D'_n(t) = -\tfrac{1}{n}g'_{n,st}(-n)exp_{n,st}(\pi i\tfrac{t}{n})exp_{n,st}(-\pi i(-n)t).$$

$$E_n(t) = \phi_n(t)D_n(t) - C_n(t)$$

$$E'_n(t) = \phi_n(t)D'_n(t) - C'_n(t)$$

$$F_n(t) = \psi_n(t)\phi_n(t)D_n(t) - \psi_n(t)C_n(t) + \phi_n(t)D'_n(t) - C'_n(t)$$

considered as $\mathfrak{C}_{n,st}$-measurable functions.

Lemma 5.21. *Discrete Fourier transform*

Let $g_{n,st}$ be $\mathfrak{C}_{n,st}$-measurable. Then, for $t \neq 0$;

$$\hat{g}_{n,st}(t) = \frac{\hat{g'}_{n,st}(t)+E_n(t)}{\psi_n(t)} = \frac{\hat{g''}_{n,st}(t)+F_n(t)}{\psi_n^2(t)}$$

Proof. We have, using Lemma 5.19(iii), that;

$$\hat{g'}_{n,st}(t) = \int_{\mathcal{R}_n} g'_{n,st}(x)exp_{n,st}(-\pi ixt)d\lambda_{n,st}(x)$$

$$= -\int_{\mathcal{R}_n} g^{sh}_{n,st}(x)exp'_{n,st}(-\pi ixt)d\lambda_{n,st}(x) + C_n(t)$$

Moreover, for $-n^2 \leq j < n^2 - 1$;

$$exp'_{n,st}(-\pi i\tfrac{j}{n}t) = n(exp_{n,st}(-\pi i\tfrac{j+1}{n}t) - exp_{n,st}(-\pi i\tfrac{j}{n}t))$$

$$= nexp_{n,st}(-\pi i\tfrac{j}{n}t)(exp_{n,st}(-\pi i\tfrac{t}{n}) - 1)$$

$$= exp_{n,st}(-\pi i\tfrac{j}{n}t)\phi_n(t).$$

Hence, noticing that $g^{sh}_{n,st}(\frac{n^2-1}{n})=0$, by Definition 5.18;

$$\hat{g'}_{n,st}(t)=-\int_{\mathcal{R}_n} g^{sh}_{n,st}(x)exp_{n,st}(-\pi ixt)\phi_n(t)d\lambda_{n,st}(x)+C_n(t)$$

$$=-\phi_n(t)\hat{g^{sh}_{n,st}}(t)+C_n(t)$$

We also have, using a change of variables, and Definition 5.18, that;

$$\hat{g^{sh}_{n,st}}(t)=\int_{\mathcal{R}_n} g^{sh}_{n,st}(x)exp_{n,st}(-\pi ixt)d\lambda_{n,st}(x)$$

$$=\int_{\frac{1-n^2}{n}}^{n} g_{n,st}(u)exp_{n,st}(-\pi i(u-\tfrac{1}{n})t)d\lambda_{n,st}(u)$$

$$=exp_{n,st}(\pi i\tfrac{t}{n})(\hat{g}_{n,st}(t)-\tfrac{1}{n}g_{n,st}(-n)exp_{n,st}(-\pi i(-n)t))$$

$$=exp_{n,st}(\pi i\tfrac{t}{n})\hat{g}_{n,st}(t)+D_n(t)$$

Therefore;

$$\hat{g'}_{n,st}(t)=-\phi_n(t)exp_{n,st}(\pi i\tfrac{t}{n})\hat{g}_{n,st}(t)-\phi_n(t)D_n(t)+C_n(t)$$

$$=\psi_n(t)\hat{g}_{n,st}(t)-E_n(t)$$

and by the same calculation;

$$\hat{g''}_{n,st}(t)=\psi_n(t)\hat{g'}_{n,st}(t)-E'_n(t)$$

$$=\psi_n(t)(\psi_n(t)\hat{g}_{n,st}(t)-E_n(t))-E'_n(t)$$

$$=\psi^2_n(t)\hat{g}_{n,st}(t)-F_n(t)$$

Rearranging, we have that, for $t\neq 0$;

$$\hat{g}_{n,st}(t)=\frac{\hat{g'}_{n,st}(t)+E_n(t)}{\psi_n(t)}=\frac{\hat{g''}_{n,st}(t)+F_n(t)}{\psi^2_n(t)}$$

as required.

□

Lemma 5.22. *If $g\in S(\mathcal{R})$, then the functions $\hat{g''}_{n,st}(t)$ and $F_n(t)$ are uniformly bounded, independently of n, for $n\geq 2$.*

Proof. Observing that;

$|D_n(t)| \leq \frac{1}{n}|g_{n,st}|(-n)$

$|D'_n(t)| \leq \frac{1}{n}|g'_{n,st}|(-n)$

$|\phi_n(t)| \leq 2n$, $|\psi_n(t)| \leq 2n$

$|C_n(t)| \leq |g_{n,st}|(\frac{n^2-1}{n}) + |g_{n,st}|(-n)$

$|C'_n(t)| \leq |g'_{n,st}|(-n)$

we obtain;

$|F_n(t)| \leq 6n|g_{n,st}|(-n) + 2n|g_{n,st}|(\frac{n^2-1}{n}) + 3|g'_{n,st}|(-n)$

$\leq 6n|g_{n,st}|(-n) + 2n|g_{n,st}|(\frac{n^2-1}{n}) + 3n|g_{n,st}|(\frac{1-n^2}{n}) + 3n|g_{n,st}|(-n)$

$= 9n|g_{n,st}|(-n) + 2n|g_{n,st}|(\frac{n^2-1}{n}) + 3n|g_{n,st}|(\frac{1-n^2}{n})$

As $g \in S(\mathcal{R})$, there exist a constant D_1, such that $|g(x)| \leq \frac{D_1}{|x|}$, $(x \neq 0)$. Then;

$|F_n(t)| \leq D_1(\frac{9n}{n} + 5\frac{n^2}{n^2-1}) \leq 16D_1$

We now calculate;

$|\hat{g''}_{n,st}|(t) = |\int_{\mathcal{R}_n} g''_{n,st}(x)exp_n(-\pi ixt)d\lambda_n(x)|$

$\leq \int_{\mathcal{R}_n} |g''_{n,st}|(x)d\lambda_n(x)$

$= \frac{1}{n}\sum_{j=-n^2}^{n^2-2} |g''_{n,st}|(\frac{j}{n})$

$= \frac{1}{n}(\sum_{j=-n^2}^{n^2-3} n|g'_{n,st}(\frac{j+1}{n}) - g'_{n,st}(\frac{j}{n})|) + D_n$

where $D_n = |g'_{n,st}|(\frac{n^2-2}{n})$.

$|\hat{g''}_{n,st}|(t) \leq (\sum_{j=-n^2}^{n^2-3} |g'_{n,st}(\frac{j+1}{n}) - g'_{n,st}(\frac{j}{n})|) + D_n$

Without loss of generality, we can assume that g is real valued, otherwise, take real and imaginary parts. Then, by the mean value theorem, for $-n^2 \leq j \leq n^2 - 3$;

$g'_{n,st}(\frac{j}{n}) = g'(\frac{j}{n} + c(j,n))$, where $0 < c(j,n) < \frac{1}{n}$

$|\hat{g''}_{n,st}|(t) \leq (\sum_{j=-n^2}^{n^2-3} |g'(\frac{j+1}{n} + c(j+1,n)) - g'(\frac{j}{n} + c(j,n))|) + D_n$

$= (\sum_{j=-n^2}^{n^2-3} |\int_{\frac{j}{n}+c(j,n)}^{\frac{j+1}{n}+c(j+1,n)} g''(x)dx|) + D_n$ (by the FTC)

$\leq (\sum_{j=-n^2}^{n^2-2} \int_{\frac{j}{n}+c(j,n)}^{\frac{j+1}{n}+c(j+1,n)} |g''|(x)dx) + D_n$

$= (\int_{-n+c(-n^2,n)}^{\frac{n^2-1}{n}+c(n^2-1,n)} |g''|(x)dx) + D_n$

$\leq (\int_{-n}^{n} |g''(x)|dx) + D_n \leq M + 2B$

where $M = ||g''||_{L^1(\mathcal{R})}$, and $B = ||g||_{C(\mathcal{R})}$.

□

Lemma 5.23. *If $g \in S(\mathcal{R})$ and $\epsilon > 0$ is standard, there exists a constant $N(\epsilon) \in \mathcal{N}_{>0}$, such that for all $n > N(\epsilon)$, for all $L, L' \in \mathcal{N}$ with $N(\epsilon) < |L| \leq |L'| \leq n$, $LL' > 0$;*

$\int_{L}^{L'} |\hat{g}_{n,st}|(t)d\lambda_n(t) < \epsilon$

Proof. We first calculate;

$|\psi_n(t)| = n|exp(\frac{\pi it}{n}) - 1|$

$= n|cos(\frac{\pi t}{n}) + isin(\frac{\pi t}{n}) - 1|$

$= n((cos(\frac{\pi t}{n}) - 1)^2 + sin(\frac{\pi t}{n})^2)^{\frac{1}{2}}$

$= n((2 - 2cos(\frac{\pi t}{n}))^{\frac{1}{2}})$

$= \sqrt{2}n(2sin^2(\frac{\pi t}{2n}))^{\frac{1}{2}}$

$= 2n|sin(\frac{\pi t}{2n})| \geq 2n(\frac{|t|}{n}) = 2|t|$ $(-n \leq t < n)$

$|\psi_n(t)|^2 \geq 4t^2$ $(-n \leq t < n)$ $(*)$

Letting W denote the bound obtained in Lemma 5.22, using Lemma 5.21, $(*)$, and, assuming, without loss of generality, that $0 \leq L \leq L'$;

$\int_L^{L'} |\hat{g}|_{n,st}(t)d\lambda_n(t)$

$\leq \int_L^n |\hat{g}|_{n,st}(t)d\lambda_n(t)$

$\leq \int_L^n \frac{W}{4t^2}d\lambda_n(t)$

$= \frac{1}{n}\sum_{j=Ln}^{n^2-1} \frac{W}{4(\frac{j}{n})^2}$

$= n\sum_{j=Ln}^{n^2-1} \frac{W}{4j^2}$

$\leq n\int_{Ln-1}^{n^2-1} \frac{W}{4x^2}dx$

$= n[\frac{-W}{4x}]_{Ln-1}^{n^2-1} = \frac{Wn}{4}(\frac{1}{Ln-1} - \frac{1}{n^2-1}) \leq \frac{W}{4}(\frac{1}{L-1} + \frac{1}{n-1}) < \epsilon$

if $min(n, L) > N(\epsilon) = \frac{W}{2\epsilon} + 1$

□

We can now show the analogous result to Theorem 5.17;

Theorem 5.24. *Let $g \in S(\mathcal{R})$, then $\hat{g}_\eta$, as given in Definition 5.15, is S-integrable on $\overline{\mathcal{R}_\eta}$, in the sense of Definition 3.26. Moreover ${}^\circ\hat{g}_\eta = st^*(\hat{g}_\infty)$, almost everywhere $L(\lambda_\eta)$, and;*

$^\circ\int_{\overline{\mathcal{R}_\eta}} \hat{g}_\eta d\lambda_\eta = \int_{\overline{\mathcal{R}_\eta}} st^*(\hat{g}_\infty)dL(\lambda_\eta) = \int_{\mathcal{R}} \hat{g}d\mu$

Proof. By Lemma 5.23;

$\mathcal{R} \models (\forall n_{(n>N(\epsilon))})(\forall L, N_{(LN\geq 0, N(\epsilon)<|L|,|N|<n)}) \int_L^N |\hat{g}_{n,st}|d\lambda_{n,st} < \epsilon$

Hence, the corresponding statement is true in ${}^*\mathcal{R}$. In particular, if η is infinite, and $\{L, N\}$ are infinite, of the same sign, belonging to $\overline{\mathcal{R}_\eta}$, we have that;

$\int_L^N |\hat{g}_\eta|d\lambda_\eta < \epsilon$

As ϵ was arbitrary we conclude that;

$\int_L^N |\hat{g}_\eta|d\lambda_\eta \simeq 0$ (*)

for all infinite $\{L, N\}$, of the same sign, in $\overline{\mathcal{R}}_\eta$. Now, using Definition 5.15 and the fact that $|exp_\eta(-\pi ixt)| \leq 1$, by transfer, we have, for $t \in \overline{\mathcal{R}}_\eta$;

$$|\hat{g}_\eta(t)| \leq \int_{\overline{\mathcal{R}_\eta}} |g_\eta(x) d\lambda_\eta = C$$

where C is finite, as, by Theorem 5.17, g_η is S-integrable. It follows that for $n \in \mathcal{N}$, the functions $\hat{g}_\eta \chi_{[-n,n]}$ are finite, in the sense of Definition 3.29. Now, proceeding as in Theorem 5.17, we obtain that $\hat{g}_\eta$ is S-integrable. Now, if $t \in \mathcal{R}_\eta$, the function $r_t(x) = g_\eta(x) exp_\eta(-\pi ixt)$ is S-integrable, by Theorem 3.30(i), as $|r_t| \leq |g_\eta|$, and g_η is S-integrable, by Theorem 5.17. Then, if t is finite, we have;

$$^\circ\hat{g}_\eta(t) = {}^\circ \int_{\overline{\mathcal{R}_\eta}} g_\eta(x) exp_\eta(-\pi ixt) d\lambda_\eta(x)$$

$$= \int_{\overline{\mathcal{R}_\eta}} {}^\circ g_\eta(x) {}^\circ exp_\eta(-\pi ixt) dL(\lambda_\eta)(x)$$

$$= \int_{x finite} st^*(g_\infty)(x) exp_\eta(-\pi i {}^\circ x {}^\circ t) dL(\lambda_\eta)(x)$$

$$= \int_{x finite} st^*(g_\infty exp_{-\pi i {}^\circ t})(x) dL(\lambda_\eta)(x)$$

$$= \int_{\mathcal{R}} g(x) exp(-\pi i {}^\circ t x) d\mu(x) = \hat{g}({}^\circ t) = st^*(\hat{g}_\infty)(t) \ (**)$$

using Definition 5.15, Theorem 3.24, Theorem 5.17, continuity of exp, see Theorem 2.25 and Theorem 5.7. Now suppose there exists $B \in L(\mathfrak{C}_\eta)$, with $L(\lambda_\eta)(B) > 0$, such that ${}^\circ\hat{g}_\eta \neq st^*(\hat{g}_\infty)$ on B. Then, by $(**)$, we can assume that $B \subset st^{-1}(\{-\infty, +\infty\})$, and $|{}^\circ\hat{g}_\eta| > 0$ on B. By the usual argument, (see Section 3), we can suppose that there exists a standard $n \in \mathcal{N}_{>0}$, with $|{}^\circ\hat{g}_\eta| > \frac{1}{n}$, on B. Then for all finite t', using Theorem 3.16;

$$^\circ \int_{|t|>t'} |\hat{g}_\eta|(t) d\lambda_\eta(t)$$

$$\geq \int_{|t|>t'} |{}^\circ\hat{g}_\eta|(t) dL(\lambda_\eta)(t) > \tfrac{1}{n} L(\lambda_\eta)(B)$$

By Lemma 2.12(i), we can find an infinite L such that;

$$\int_{|t|>L} |\hat{g}_\eta|(t) d\lambda_\eta)(t) > \tfrac{1}{2n} L(\lambda_\eta)(B)$$

This contradicts $(*)$. Hence ${}^\circ\hat{g}_\eta = st^*(\hat{g}_\infty)$ a.e $L(\lambda_\eta)$. the rest of the proof is the same as Theorem 5.17. □

Finally, we have;

Theorem 5.25. *For $g \in S(\mathcal{R})$, the Fourier Inversion Theorem holds and admits a non standard proof.*

Proof. By Lemma 5.16, we have that;

$$g_\eta(x) = \frac{1}{2}\int_{\overline{\mathcal{R}_\eta}} \hat{g}_\eta(t) exp_\eta(\pi ixt) d\lambda_\eta(t) \ (*)$$

for $x \in \overline{\mathcal{R}_\eta}$. As in Theorem 5.24, the function $s_x(t) = \hat{g}_\eta(t)exp_\eta(\pi ixt)$ is S-integrable, because, by the same theorem, $\hat{g}_\eta$ is S-integrable. We now argue as before, and use the result that ${}^\circ g_\eta = st^*(\hat{g}_\infty)$, a.e $L(\lambda_\eta)$. We have, if x is standard, taking standard parts in $(*)$;

$$g(x) = {}^\circ g_\eta(x) = \frac{1}{2}\int_{\overline{\mathcal{R}_\eta}} {}^\circ\hat{g}_\eta(t) {}^\circ exp_\eta(\pi ixt) dL(\lambda_\eta)(t)$$

$$= \frac{1}{2}\int_{t finite} st^*(\hat{g}_\infty)(t) exp_\eta(\pi ix {}^\circ t) dL(\lambda_\eta)(t)$$

$$= \frac{1}{2}\int_{t finite} st^*(\hat{g}_\infty exp_{\pi ix})(t) dL(\lambda_\eta)(t)$$

$$= \frac{1}{2}\int_{\mathcal{R}} \hat{g}(t) exp(\pi ixt) d\mu(t)$$

as required.

□

Remarks 5.26. *It would be interesting to give a nonstandard proof of the Inversion theorem in greater generality, that is without the assumption that $g \in S(\mathcal{R})$. Further developments might include whether one can obtain Plancherel formulae for Lie groups from finite groups, using nonstandard methods, see* [40] *for a statement of the Plancherel Theorem over $\mathcal{R}$.*

6. The Ergodic Theorem

There are many versions of the ergodic theorem, but the one we will prove in this section, using nonstandard analysis, is the following;

Theorem 6.1. *Ergodic Theorem*

Let $(\Omega, \mathfrak{C}, \mu)$ be a probability space, and let T be a measure preserving transformation, then, if $g \in L^1(\Omega, \mathfrak{C}, \mu)$;

$\diamond g(\omega) = lim_{n\rightarrow\infty}\frac{1}{n}\sum_{i=0}^{n-1} g(T^i\omega)$

exists for almost all $\omega \in \Omega$*, with respect to* μ*, and,* $\diamond g \in L^1(\Omega, \mathfrak{C}, \mu)$*, with;*

$\int_\Omega \diamond g d\mu = \int_\Omega g d\mu$

Remarks 6.2. *There are a number of good standard proofs of this result. A particular good reference is* [31]. *However, the reader should be aware that it is assumed there that* $\mathfrak{C}$ *is complete and* T *is*invertible, *in the sense that* T *is one-one and onto, and both* T *and* T^{-1} *are measurable. A m.p.t is then required to satisfy* $\mu(C) = \mu(T^{-1}C)$ *for all* $C \in \mathfrak{C}$*. We will not require these assumption in the proofs of this section, in the sense that we only require a m.p.t to be a measurable* T *with* $\mu(C) = \mu(T^{-1}C)$ *for all* $C \in \mathfrak{C}$*. In* [31], *a seemingly stonger result is shown, (under the above assumptions), namely that if* $C \in \mathfrak{C}$*, with* $T^{-1}(C) = C$*, then;*

$\int_C \diamond g d\mu = \int_C g d\mu$ $(*)$

from which it easily follows that if $\mathfrak{C}'$ *is the sub* σ*-algebra of all* T*-invariant sets, where a set* C *is* T *invariant in* [31], *if* $T^{-1}C = C$ *a.e* $d\mu$*, then* $\diamond g = E(g|\mathfrak{C}')$*,* $(**)$*. In the particular case when* T *is ergodic, that is every* T *invariant set has measure* 0 *or* 1*, we obtain the well known result that* $\diamond g = E(g)$ *a.e* $d\mu$*,* $(***)$*. However, this result* $(*)$ *follows easily from our Theorem 6.1. as we can, wlog, assume that* $\mu(C) > 0$*, and then restrict and rescale the measure. Of course, we even obtain a slight strengthening of* $(*)$*, by our weaker assumption on a m.p.t, and obtain similar strengthenings of* $(**)$ *and* $(***)$*. (It is not necessary to restrict attention to real valued functions, in the statement of the theorem, the complex version follows immediately from the real case).*

As usual, we work in an $\aleph_1$-saturated model. Let $k \in {}^*\mathcal{N}_{>0}$ be infinite, and let $K = \{x \in {}^*\mathcal{N} : 0 \leq x < k\}$. We let $\mathfrak{K}$ be the algebra of all internal subsets of K. Observe that as K is hyperfinite, $\mathfrak{K}$ is a hyperfinite ${}^*\sigma$-algebra. We let ν denote the counting measure, defined by setting $\nu(A) = \frac{Card(A)}{k}$, for $A \in \mathfrak{K}$. As in Section 3, we let $P = {}^\circ\nu$. By Theorem 3.4, and remarks before Lemma 3.15, P extends uniquely to the completion $\mathfrak{B}$ of the σ-algebra, $\sigma(\mathfrak{K})$, generated by $\mathfrak{K}$. It is clear that $(K, \mathfrak{B}, P)$ is a probability space, it is also the Loeb space associated to $(K, \mathfrak{K}, \nu)$. We let $\phi : K \rightarrow K$ denote the map defined by;

$\phi(x) = x + 1$, if $0 \leq x < k - 1$

$\phi(x) = 0$, if $x = k - 1$

Clearly, ϕ is invertible, internal, preserves the counting measure ν, and $\phi^{-1}(\sigma(\mathfrak{K})) = \sigma(\mathfrak{K})$. Then $P \circ \phi^{-1}$ defines a measure on $(K, \sigma(\mathfrak{K}), P)$, extending ν. By Theorem 3.4(ii), it agrees with P. By definition of the completion, $P \circ \phi^{-1}$ agrees with P on $(K, \mathfrak{B}, P)$, so ϕ, and similarly ϕ^{-1} are m.p.t's. We will first prove the following;

Theorem 6.3. *The ergodic theorem, as stated in Theorem 6.1, holds for* $(K, \mathfrak{B}, P, \phi)$.

Proof. Let $g \in L^1(K, \mathfrak{B}, P)$, without loss of generality, we can assume that $g \geq 0$. For $x \in K$, we let;

$\overline{g}(x) = limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} g(\phi^i x)$

$\underline{g}(x) = liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} g(\phi^i x)$

In order to prove the theorem, it is sufficient to show that $\overline{g}$ is integrable and;

$\int_K \overline{g} dP \leq \int_K g dP \leq \int_K \underline{g} dP$ (†)

Then, as $\underline{g} \leq \overline{g}$, we must have equality in (†), so $\underline{g} = \overline{g}$ a.e dP, that is $\diamond g$ exists a.e dP, and;

$\int_K \diamond g dP = \int_K g dP$

as required.

Now let $M \in \mathcal{N}_{>0}$, then, as $\overline{g}$ is $\mathfrak{B}$-measurable, see [37], $min(\overline{g}, M)$ is integrable with respect to P. Let $\epsilon > 0$ be standard, then we can apply Theorem 3.36, Definition 3.9 and Remarks 3.10, to obtain internal functions $F, G : K \rightarrow^* \mathcal{R}$, with $g \leq F$ and $G \leq min(\overline{g}, M)$, such that;

$| \int_A g dP - \frac{1}{k} {}^* \sum_{x \in A} F(x) | < \epsilon$

$|\int_A min(\overline{g}, M)dP - \frac{1}{k}{}^*\sum_{x\in A} G(x)| < \epsilon$, for all internal $A \subset K$, $(\dagger\dagger)$.

Now observe that $\overline{g}$ is ϕ-invariant,([12]). Fixing $x \in K$, by the definition of $\overline{g}$, we can find $n \in \mathcal{N}_{>0}$ such that;

$min(\overline{g}(x), M) \leq \frac{1}{n}\sum_{i=0}^{n-1} g(\phi^i x) + \epsilon$ $(*)$

Then, if $0 \leq m \leq n-1$, we have;

$G(\phi^m x) \leq min(\overline{g}(\phi^m x), M)$, by definition of G

$= min(\overline{g}(x), M)$, by ϕ invariance of $\overline{g}$

$\leq \frac{1}{n}\sum_{i=0}^{n-1} g(\phi^i x) + \epsilon$, by $(*)$

$\leq \frac{1}{n}\sum_{i=0}^{n-1} F(\phi^i x) + \epsilon$, by definition of F

Therefore,

$\sum_{i=0}^{n-1} G(\phi^i x) \leq n(\frac{1}{n}\sum_{i=0}^{n-1} F(\phi^i x) + \epsilon) = \sum_{i=0}^{n-1} F(\phi^i x) + n\epsilon$ $(**)$

Now let $S_G : [1, k) \times K \rightarrow {}^*\mathcal{R}$ be defined by;

$S_G(n, x) = {}^*\sum_{i=0}^{n-1} G(\phi^i x)$

and, similarly, define S_F. By Definition 2.19, and using the facts that K is $*$-finite, and G, F are internal, S_G and S_F are internal. Then, the relation $(**)$ becomes the internal relation on $[1, k) \times K$, given by $R(n, x)$ iff $S_G(n, x) \leq S_F(n, x) + n\epsilon$. Using the fact above, that the fibres of R over K are non-empty, by transfer of the corresponding standard result, we can find an internal function $T : K \rightarrow [1, k)$, which assigns to $x \in K$, the least $n \in [1, k)$, for which $(**)$ holds. Moreover, as we have observed in $(*)$, $T(x)$ is standard, for all $x \in K$. By Lemma 3.11, $r = max_{x\in K}T(x)$ exists and is standard. Now, define T_j hyper

[12]There is a probably a proof of this result in the literature, but we supply one here. Fix $x \in K$. Let $A_m = \frac{1}{m}\sum_{i=0}^{m-1} g(\phi^i x)$ and let $B_m = \frac{1}{m}\sum_{i=0}^{m-1} g(\phi^{i+1} x)$. Then a simple calculation shows that $\frac{mB_m + g(x)}{m+1} = A_{m+1}$. Hence, $|B_m - A_{m+1}| = |\frac{A_{m+1} - g(x)}{m}|$, $(*)$. Suppose that $\overline{g}(x) = t < \infty$, $(**)$, (the case when $\overline{g}(x) = \infty$ is similar), and $\overline{g}(\phi x) < t$, $(***)$, (the case $\overline{g}(\phi x) > t$ is again similar). Then, by $(***)$, there exists $\delta > 0$, such that, for $m \geq m_0$, $B_m < t - \delta$. By $(*)$ and $(**)$, we can find $m_1 \geq m_0$, such that $|B_m - A_{m+1}| < \frac{\delta}{2}$, for $m \geq m_1$. Again, by $(*)$, we can find $m_2 \geq m_1 \geq m_0$, such that $A_{m_2+1} > t - \frac{\delta}{2}$. This clearly gives a contradiction.

inductively by;

$T_0 = 0$ and $T_j = T_{j-1} + T(T_{j-1})$

and let J be the first j such that $k - r \leq T_j < k$.([13])

Observe that T_j defines an internal partition of the interval $[0, T_{J-1}] \subset [0, k)$, into $J - 1$ blocks of step size $T_j - T_{j-1} = T((T_{j-1})$. Hence, we can write;

$\frac{1}{k}{}^*\sum_{x=0}^{T_J-1} G(x) = \frac{1}{k}{}^*\sum_{j=0}^{J-1}{}^*\sum_{i=0}^{T(T_j)-1} G(\phi^i T_j)$

$\leq \frac{1}{k}{}^*\sum_{j=0}^{J-1}{}^*\sum_{i=0}^{T(T_j)-1} F(\phi^i T_j) + T(T_j)\epsilon$,by definition of T and $(**)$.

Now we can rearrange this last sum as;

$\frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \frac{\epsilon}{k}{}^*\sum_{j=0}^{J-1} T(T_j)$

$= \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \frac{T_J\epsilon}{k}$

$< \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \epsilon$

using the facts that ${}^*\sum_{j=0}^{J-1} T(T_j) = {}^*\sum_{j=0}^{J-1}(T_{j+1} - T_j) = T_J$, and $T_J < k$. Therefore, we have that;

$\frac{1}{k}{}^*\sum_{x=0}^{T_J-1} G(x) < \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + \epsilon \; (***)$

Now, observing that $\nu([T_J, k)) \leq \frac{r}{k} \simeq 0$, as r is standard, we have $P([T_J, k)) = 0$. Hence, using $(\dagger\dagger)$, $(***)$;

$\int_X min(\overline{g}, M)dP = \int_{[0,T_J)} min(\overline{g}, M)dP < \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} G(x) + \epsilon$

[13]This perhaps requires some explanation. Define $I = \{m \in {}^*\mathcal{N}_{>0} : \exists! S(dom(S) = [0, m] \wedge S(0) = 0 \wedge (\forall 1 \leq j \leq m) S(j) = S(j-1) + T(S(j-1)_{modk}))\}$, $(*)$, then it is easy to see that I is internal, $I(1)$ holds, and $I(m)$ implies $I(m+1)$. Applying Lemma 2.12, $I = {}^*\mathcal{N}_{>0}$. Hence there exists an internal function f, defined on ${}^*\mathcal{N}_{>0}$, such that $f(m)$ is the unique S satisfying $(*)$. We can then define $T_j = f(j)(j)$, and clearly $T_j - T_{j-1} \leq r$. Let $V = \{j \in {}^*\mathcal{N}_{>0} : T_j < k\}$. Then, as $T \geq 1$, V is the interval $[1, t]$ for some infinite $t < k$. Then $k - r \leq T_t < k$, otherwise $T_{t+1} < k$. Then $U = \{j \in {}^*\mathcal{N}_{>0} : k - r \leq T_j < k\}$ is internal and non empty. Therefore, by transfer, it contains a first element J.

$< \frac{1}{k}{}^*\sum_{x=0}^{T_J-1} F(x) + 2\epsilon < \int_{[0,TJ)} gdP + 3\epsilon = \int_X gdP + 3\epsilon$

Now, letting $M \rightarrow \infty$ and $\epsilon \rightarrow 0$, we can apply the MCT, to obtain;

$\int_X \overline{g}dP \leq \int_X gdP$

As g is integrable with respect to P, so is $\overline{g}$, and a similar argument to the above demonstrates that $\int_X gdP \leq \int_X \underline{g}dP$. Therefore, (†) is shown and the theorem is proved.

□

We now generalise Theorem 6.3, to obtain Theorem 6.1. We let $\mathcal{P}$ consist of spaces of the form $(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda, \sigma)$, where $\mathfrak{D}$ is the Borel field on $\mathcal{R}^{\mathcal{N}}$, σ is the left shift on $\mathcal{R}^{\mathcal{N}}$, and λ is a shift invariant probability measure. Note that σ is not invertible, but we require that $\lambda = \sigma_*\lambda$, so σ is a m.p.t, with respect to λ. Similarly, we let $\mathcal{Q}$ consist of spaces of the form $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$, where $\mathfrak{E}$ is the Borel field on $[0,1]^{\mathcal{N}}$, σ is again the left shift, and ρ is a shift invariant probability measure.

We first require the following simple lemma;

Lemma 6.4. *Theorem 6.1 is true iff the Ergodic Theorem holds for all spaces in $\mathcal{P}$.*

Proof. One direction is obvious. For the other direction, let $(\Omega, \mathfrak{C}, \mu, T)$ and $g \in L^1(\Omega, \mathfrak{C}, \mu)$ be given. Define a map $\tau : \Omega \rightarrow \mathcal{R}^{\mathcal{N}}$ by $\tau(\omega)(n) = g(T^n\omega)$. Clearly, as g is measurable with respect to $\mathfrak{C}$ and T is a m.p.t, using the definition of the Borel field on $\mathcal{R}^m$, for finite m, we have that for a cylinder set $U \in \mathfrak{D}$, $\tau^{-1}(U) \in \mathfrak{C}$. By the definition of the Borel field on $\mathcal{R}^{\mathcal{N}}$, $\tau^{-1}(\mathfrak{D}) \subset \mathfrak{C}$, ([14]). Let λ be the probability measure $\tau_*\mu$. Then λ is σ invariant, as clearly, using the fact that T is a m.p.t, $\lambda = \sigma_*\lambda$ on the cylinder sets in $\mathfrak{D}$. Using the definition of the Borel field and Caratheodory's Theorem, we obtain that $\lambda = \sigma_*\lambda$. Let $\pi : \mathcal{R}^{\mathcal{N}} \rightarrow \mathcal{R}$ be the projection onto the $0'th$ coordinate. Then $g = \pi \circ \tau$, and, so $\pi \in L^1(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda)$ by the change of variables formula, ([15]). Moreover, $g(T^i\omega) = \pi(\sigma^i\tau(\omega))$, so applying the Ergodic Theorem for $(\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda, \sigma)$, with the change of variables formula, we have that

[14] As $\{V \in \mathfrak{D} : \tau^{-1}(V) \in \mathfrak{C}\}$ is a σ-algebra containing the cylinder sets.

[15] This states that if $\tau : (X_1, \mathfrak{C}_1, \mu_1) \rightarrow (X_2, \mathfrak{C}_2, \mu_2)$ is measurable and measure preserving, so $\mu_2 = \tau_*\mu_1$, then a function $\theta \in L^1(X_2, \mathfrak{C}_2, \mu_2)$ iff $\tau^*\theta \in L^1(X_1, \mathfrak{C}_1, \mu_1)$ and $\int_C \theta d\tau_*\mu_1 = \int_{\tau^{-1}(C)} \tau^*\theta d\mu_1$.

$\diamond g$ exists and $\diamond g = \diamond\pi \circ \tau$ a.e $d\mu$, and $\int_\Omega \diamond g d\mu = \int_\Omega (\diamond\pi \circ \tau) d\mu = \int_{\mathcal{R}^\mathcal{N}} \diamond\pi d\lambda = \int_{\mathcal{R}^\mathcal{N}} \pi d\lambda = \int_\Omega g d\mu$ as required.

□

We make the following definition;

Definition 6.5. *We say that* $(\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda, \sigma) \in \mathcal{P}$ *is a factor of* $(K, \mathfrak{B}, P, \phi)$ *if there exists;*

$\Gamma : (K, \mathfrak{B}, P) \rightarrow (\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda)$

which is measurable and measure preserving, such that;

$\Gamma(\phi x) = \sigma(\Gamma x)$ *a.e* $(x \in K)$ dP.

We make the same definition if $([0,1]^\mathcal{N}, \mathfrak{E}, \rho, \sigma) \in \mathcal{Q}$.

Lemma 6.6. *Suppose that* $(\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda, \sigma) \in \mathcal{P}$ *is a factor of* $(K, \mathfrak{B}, P, \phi)$, *then, if the Ergodic Theorem holds for* $(K, \mathfrak{B}, P, \phi)$, *it holds for* $(\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda, \sigma)$.

Proof. The proof is similar to Lemma 6.4. If $h \in L^1(\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda)$, then, by change of variables, $\Gamma^* h \in L^1(K, \mathfrak{B}, P)$. Applying the Ergodic Theorem for $(K, \mathfrak{B}, P, \phi)$ and the definition of a factor, we have that $\diamond\Gamma^* h$ exists and $\diamond\Gamma^* h = \Gamma^* \diamond h$, a.e d$P$, $(*)$. So $\diamond h$ exists a.e dλ, and, again, by change of variables, $(*)$, and the Ergodic theorem for $(K, \mathfrak{B}, P, \phi)$;

$$\int_{\mathcal{R}^\mathcal{N}} \diamond h d\lambda = \int_K \Gamma^*(\diamond h) dP = \int_K \diamond(\Gamma^* h) dP = \int_K (\Gamma^* h) dP = \int_{\mathcal{R}^\mathcal{N}} h d\lambda$$

□

We now claim the following;

Lemma 6.7. *Every space in* $\mathcal{P}$ *is isomorphic, in the sense of dynamical systems, ([16]), to a space in* $\mathcal{Q}$.

Proof. There exists an isomorphism, in the sense of measure spaces, $\Phi : (\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda) \rightarrow ([0,1], \mathfrak{E}', \rho')$, where $\mathfrak{E}'$ is the Borel field and ρ' is a probability measure, see [31], Theorem 1.4.4. Now define $r : \mathcal{R}^\mathcal{N} \rightarrow [0,1]^\mathcal{N}$ by $r(\omega)(n) = \Phi(\sigma^n \omega)$. Again, using the argument above and the

[16] By which I mean there exists measurable and measure preserving maps $r : (\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda) \rightarrow ([0,1]^\mathcal{N}, \mathfrak{E}, \rho)$ and $s : ([0,1]^\mathcal{N}, \mathfrak{E}, \rho) \rightarrow (\mathcal{R}^\mathcal{N}, \mathfrak{D}, \lambda)$ such that $s \circ r = Id$ and $r \circ \sigma = \sigma \circ r$ a.e dλ, $r \circ s = Id$ and $s \circ \sigma = \sigma \circ s$ a.e dρ

fact that Φ and σ are measurable, $r^{-1}(\mathfrak{E}) \subset \mathfrak{D}$, where is the Borel field on $[0,1]^{\mathcal{N}}$. Let ρ be the probability measure $r_*\lambda$, so $r : (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda) \rightarrow ([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho)$ is also measure preserving. We have that $r(\sigma\omega)(n) = \Phi(\sigma^{n+1}\omega) = (r\omega)(n+1) = \sigma(r\omega)(n)$, so $r \circ \sigma = \sigma \circ r$, for all $\omega \in \mathcal{R}^{\mathcal{N}}$. This also shows that ρ is σ invariant, as λ is σ invariant. Hence, $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ belongs to $\mathcal{Q}$. Define $s : ([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho) \rightarrow (\mathcal{R}^{\mathcal{N}}, \mathfrak{D}, \lambda)$, by, $s(\omega') = \Phi^{-1}(\pi(\omega'))$, where again π is the 0'th coordinate projection, clearly s is measurable. Then $(s \circ r)(\omega) = \Phi^{-1} \circ \pi \circ r(\omega)$, and $\pi \circ r(\omega) = r(\omega)(0) = \Phi(\omega)$, so $(s \circ r) = Id$ a.e, and, similarly $r \circ \sigma = \sigma \circ r$ a.e $d\lambda$. This clearly shows that s is measure preserving, and that $(r \circ s) = Id$, $s \circ \sigma = \sigma \circ s$,$(*)$, hold, restricted to $r(U)$, where $\lambda(U) = 1$. As, by definition, $\rho(\lambda(U)) = 1$, and the conditions in $(*)$ are measurable, we obtain the result. (Note that the map s need not be invertible in the ordinary sense.) □

We now make the following;

Definition 6.8. *Let* $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ *belong to* $\mathcal{Q}$*, then we say that* α *is typical for* ρ *if;*

$$lim_{n\rightarrow\infty} \sum_{i=0}^{n-1} g(\sigma^i \alpha) = \int_{[0,1]^{\mathcal{N}}} g d\rho$$

for any $g \in C([0,1]^{\mathcal{N}})$.

We now show;

Theorem 6.9. *Let* $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ *belong to* $\mathcal{Q}$*, possessing a typical element* α*. Then* $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma)$ *is a factor of* $(K, \mathfrak{B}, P, \phi)$ *in the sense of Definition 6.5.*

Proof. Define $\Gamma : K \rightarrow [0,1]^{\mathcal{N}}$ by $\Gamma(x) = {}^\circ(\sigma^x \alpha)$, ([17]). Now suppose that $g \in C([0,1]^{\mathcal{N}})$, so, as $[0,1]^{\mathcal{N}}$ is compact, g is bounded,$(*)$, then;

[17]Here, $(\sigma^x \alpha) = {}^*H(x)$ for the internal function ${}^*H : {}^*\mathcal{N} \rightarrow {}^*([0,1]^{\mathcal{N}}) = ({}^*[0,1])^{{}^*\mathcal{N}}$, obtained by transferring the standard function $H : \mathcal{N} \rightarrow [0,1]^{\mathcal{N}}$, defined by $H(n) = \sigma^n(\alpha)$. Observe that $[0,1]^{\mathcal{N}}$ is compact and Haussdorff in the product topology, so, by Theorem 2.34, there exists a unique standard part mapping ${}^\circ : {}^*([0,1]^{\mathcal{N}}) \rightarrow [0,1]^{\mathcal{N}}$. In fact, see [36], this mapping is defined by setting ${}^\circ s = ({}^\circ s(n))_{n\in\mathcal{N}}$ where $s : {}^*\mathcal{N} \rightarrow^* [0,1]$ is internal.

${}^{\circ}g(\sigma^x\alpha) = g(\Gamma(x))$ for all $x \in K$, $(**)$ ([18]).

This implies that Γ is measurable, as if B is an open set for the product topology on $[0,1]^{\mathcal{N}}$, then, taking g to be a continuous function with support B, $\Gamma^* g$ is measurable with respect to P, by Theorem 3.8 (Lemma 3.15) This clearly implies that $\Gamma^{-1}(B)$ is measurable. By previous arguments, we obtain the result. Moreover;

$\int_{[0,1]^{\mathcal{N}}} g d\rho$

$= lim_{n\rightarrow\infty}\frac{1}{n}\sum_{i=0}^{n-1} g(\sigma^i\alpha)$, (by definition of a typical element α)

$=^{\circ} (\frac{1}{k}{}^{*}\sum_{x=0}^{k-1} g(\sigma^x\alpha))$, ([19]).

$= {}^{\circ}\int_K g(\sigma^x\alpha)d\nu$ (using Definition 3.9 and Remarks 3.10)

$= \int_K g(\Gamma(x))dP$, (using $(*)$, $(**)$ and Theorem 3.12 (Lemma 3.15))

$(***)$

The result of $(***)$ implies that Γ is measure preserving. The probability measure $\Gamma_* P$ defines a bounded linear functional on $C([0,1]^{\mathcal{N}})$, which agrees with ρ. Using the fact that $[0,1]^{\mathcal{N}}$ is a compact Hausdorff space, and $\rho, \Gamma_* P$ are regular, see [37] Theorem 2.18, ([20]), we can apply the uniqueness part of the Riesz Representation Theorem, see [37] Theorem 6.19, to conclude that $\Gamma_* P = \rho$, we will discuss this further below. Now, as σ is continuous with respect to $\mathfrak{C}$, ([21]),;

---

[18] I have also denoted by g, the transfer of g to ${}^*C({}^*([0,1]^{\mathcal{N}}))$. Observe that $\sigma^x(\alpha) \simeq \Gamma(x)$ by definition of Γ, it is then straightforward to adapt Theorem 2.25, using the fact that g is continuous, to show that $g(\sigma^x\alpha) \simeq g(\Gamma(x))$.

[19] Observe that $s(n) = \frac{1}{n}\sum_{i=0}^{n-1} g(\sigma^i\alpha)$ is a standard sequence, with limit $s = \int_{[0,1]^{\mathcal{N}}} g d\rho$. By Theorem 2.22, using the fact that k is infinite, $s \simeq s(k)$. Using Definition 2.19, it is clear that $s(k)$ is the hyperfinite sum $\frac{1}{k}{}^{*}\sum_{x=0}^{k-1} g(\sigma^x\alpha)$

[20] It is easy to see that $[0,1]^{\mathcal{N}}$ is σ-compact. This follows from the fact that finite intersections of cylinder sets form a basis for the topology on $[0,1]^{\mathcal{N}}$. Any open set in U in $[0,1]^m$ is a countable union of closed sets, as every $x \in U$ lies inside a closed box B with rational corners, such that $B \subset U$. Hence, any cylinder set is a countable union of such closed sets $\pi_m^{-1}(B)$.

[21] Again I have denoted by σ the transfer of the standard shift σ to ${}^*([0,1]^{\mathcal{N}})$. The fact that $\sigma(\sigma^x\alpha) = \sigma^{x+1}(\alpha)$ follows immediately by transferring the standard fact that $\sigma(\sigma^n(\alpha)) = \sigma^{n+1}(\alpha)$ for $n \in \mathcal{N}$.

$\sigma(\Gamma x) = \sigma(^\circ(\sigma^x \alpha) =^\circ (\sigma(\sigma^x \alpha)) =^\circ (\sigma^{x+1}\alpha) = \Gamma(x+1) = \Gamma(\phi(x))$

except for $x = k - 1$, so a.e dP. Hence, the result follows.

□

We now address the problem of finding a typical element for a space $([0,1]^{\mathcal{N}}, \mathfrak{E}, \rho, \sigma) \in \mathcal{Q}$. By Theorem 6.3, Lemma 6.4, Lemma 6.6, Lemma 6.7 and Theorem 6.9, we then obtain the Ergodic Theorem 6.1. The proof of this result does *not* require the Ergodic Theorem, and is originally due to de Ville, see [21].

Definition 6.10. *We say that a sequence of measures $(\rho_n)_{n\in\mathcal{N}}$ converges weakly to ρ if, for all $g \in C([0,1]^{\mathcal{N}})$;*

$lim_{n\rightarrow\infty}(\int_{[0,1]^{\mathcal{N}}} g d\rho_n) = \int_{[0,1]^{\mathcal{N}}} g d\rho.$

We require the following lemma;

Lemma 6.11. *Let $(\alpha_n)_{n\in\mathcal{N}}$ be a sequence of periodic, with respect to σ, elements in $[0,1]^{\mathcal{N}}$, such that the sequence of probability measures $(\rho_{\alpha_n})_{n\in\mathcal{N}}$ converges weakly to ρ, where;*

$\rho_{\alpha_n} = \frac{1}{c_n}(\delta_{\alpha_n} + \delta_{\sigma\alpha_n} + \ldots + \delta_{\sigma^{c_n-1}\alpha_n})$

δ_{α_n} denotes the probability measure supported on α_n and c_n denotes the period of α_n. Then there exists a sequence $(r_n)_{n\in\mathcal{N}}$ of positive integers, such that if $(T_n)_{n\in\mathcal{N}}$ is defined by $T_0 = 0$ and $T_{n+1} - T_n = c_n r_n$, the element $\alpha \in [0,1]^{\mathcal{N}}$, defined by $\alpha(m) = \alpha_n(m - T_n)$, for $T_n \leq m < T_{n+1}$, is typical for ρ.

Proof. The proof is intuitively clear, but hard to write down rigorously. As ρ_{α_n} converges weakly to ρ, we have that;

$lim_{n\rightarrow\infty}(\int_X f d\rho_{\alpha_n}) = \int_X f d\rho$

By definition of ρ_{α_n};

$\int_X f d\rho_{\alpha_n} = \frac{1}{c_n}(f(\alpha_n) + \ldots + f(\sigma^{c_n-1}\alpha_n)$

So it is sufficient to prove that;

$lim_{n\rightarrow\infty}\frac{1}{n}\sum_{i=0}^{n-1} f(\sigma^i\alpha) = lim_{n\rightarrow\infty}\frac{1}{c_n}(f(\alpha_n) + \ldots + f(\sigma^{c_n-1}\alpha_n))$ (*)

We first claim that, if $f \in C([0,1]^{\mathcal{N}})$, there exists an increasing sequence $\{m_n\}_{n\in\mathcal{N}}$ of positive integers, such that if $b, c \in [0,1]^{\mathcal{N}}$, and agree up to the m_n'th coordinate, then $|f(b) - f(c)| < \frac{1}{n}$, $(**)$. In order to see this, for $x \in [0,1]^{\mathcal{N}}$, let $U_x = \{y : |f(x) - f(y)| < \frac{1}{2n}\}$. As f is continuous, U_x is open in the Borel field, hence there exists $V_x \subset U_x$, containing x, of the form $\pi^{-1}(W_x)$, where $W_x \subset \mathcal{R}^{n_x}$ is open, and π is the projection onto the first n_x coordinates. Then, if $y, z \in U_x$, $|f(y) - f(z)| \leq |f(y) - f(x)| + |f(z) - f(x)| < \frac{1}{n}$. The sets $\{V_x : x \in X\}$ form an open cover of $[0,1]^{\mathcal{N}}$, which is compact in the product topology. Hence, there exists a finite subcover $V_{x_1} \cup \ldots \cup V_{x_r}$. We can choose m_n such that each V_{x_j} is of the form $\pi^{-1}(W_{x_j})$, for $W_{x_j} \subset \mathcal{R}^{m_n}$. Then, if b and c agree up to the m_n'th coordinate, we have that $b \in V_{x_j}$ iff $c \in V_{x_j}$, so $|f(b) - f(c)| < \frac{1}{n}$, showing $(**)$. Now let $\{g_n\}_{n\in\mathcal{N}}$ be any increasing sequence of positive integers, such that if $Q_n = sup\{|f(b) - f(c)| : \pi_{g_n}(b) = \pi_{g_n}(c)\}$, then $\{Q_n\}_{n\in\mathcal{N}}$ is decreasing and $lim_{n\rightarrow\infty} Q_n = 0$. Clearly such a sequence exists by $(**)$. Without loss of generality, we can choose $\{g_n\}_{n\in\mathcal{N}}$, such that the periods $c_n | g_n$, $(\sharp)$. Now choose $\{T_i\}_{i\in\mathcal{N}}$ as follows;

(i). $T_{i+1} \geq 2^i T_i$

(ii). $g_i | T_{i+1} - T_i$ (so $c_i | T_{i+1} - T_i$)

(iii). $C_i = \frac{T_{i+1} - T_i}{g_i} \gneq C_{i-1} = \frac{T_i - T_{i-1}}{g_{i-1}}$ $(i \geq 1)$.

(iv). $T_i \geq 2^i c_i$ $(i \geq 1)$.

We now claim there exists a decreasing sequence $\{b_n\}_{n\in\mathcal{N}_{>0}}$ of positive reals, such that;

$$|\frac{1}{T_n} \sum_{i=0}^{T_n - 1} f(\sigma^i \alpha) - t_n| \leq b_n \ (***)$$

where $lim_{n\rightarrow\infty} b_n = 0$, and $t_n = \frac{1}{c_n}(f(\alpha_n) + \ldots + f(\sigma^{c_n - 1}\alpha_n))$, for $n \geq 1$. For ease of notation, we let;

$$A_n = \frac{1}{n} \sum_{i=0}^{n-1} f(\sigma^i \alpha)$$

$$A_{m,n} = \frac{1}{n-m} \sum_{i=m}^{n-1} f(\sigma^i \alpha)$$

Recall the law of weighted averages, $A_n = \frac{mA_m+(n-m)A_{m,n}}{n}$. We first estimate $|A_{T_n} - A_{T_{n-1},T_n}|$. We have;

$A_{T_n} = \frac{T_{n-1}A_{T_{n-1}}+(T_n-T_{n-1})A_{T_{n-1},T_n}}{T_n}$

$|A_{T_n} - A_{T_{n-1},T_n}|$

$= |\frac{T_{n-1}}{T_n}A_{T_{n-1}} + \frac{T_n-T_{n-1}}{T_n}A_{T_{n-1},T_n} - A_{T_{n-1},T_n}|$

$\leq \frac{|A_{T_{n-1}}|}{2^{n-1}} + \frac{|A_{T_{n-1},T_n}|}{2^{n-1}}$ by (i)

$\leq \frac{M}{2^{n-2}}$, where $|f| \leq M$, (A)

We now estimate the average A_{T_{n-1},T_n}. The idea is to divide the interval between T_{n-1} and T_n into C_{n-1} blocks of length g_{n-1}, where the period $c_{n-1}|g_{n-1}$, using $(\sharp)$ and (ii). We estimate $|A_{T_{n-1},T_n} - A_{T_{n-1},T_n-g_n}|$;

$A_{T_{n-1},T_n} = \frac{C_{n-1}-1}{C_{n-1}}A_{T_{n-1},T_n-g_{n-1}} + \frac{1}{C_{n-1}}A_{T_n-g_{n-1},T_n}$

$|A_{T_{n-1},T_n} - A_{T_{n-1},T_n-g_{n-1}}|$

$= |\frac{A_{T_n-g_{n-1},T_n}}{C_{n-1}} - \frac{A_{T_{n-1},T_n-g_{n-1}}}{C_{n-1}}| \leq \frac{2M}{C_{n-1}}$ (B)

We now let;

$B_{T_{n-1},m} = \frac{1}{m-T_{n-1}}\sum_{i=0}^{m-T_{n-1}-1} f(\sigma^i\alpha_{n-1})$, for $m \leq n$.

We estimate $|A_{T_{n-1},T_n-g_{n-1}} - B_{T_{n-1},T_n-g_n}|$. We have that $\sigma^{T_{n-1}+i}\alpha$ and $\sigma^i\alpha_{n-1}$ agree up to the g_{n-1}'th coordinate, for $0 \leq i < T_n - T_{n-1} - g_{n-1}$. Therefore, for such i, $|f(\sigma^i\alpha_{n-1}) - f(\sigma^{T_{n-1}+i}\alpha)| \leq Q_{n-1}$, and so;

$|A_{T_{n-1},T_n-g_{n-1}} - B_{T_{n-1},T_{n-1}-g_{n-1}}| \leq Q_{n-1}$ (C)

Now, by the same argument as in (B);

$|B_{T_{n-1},T_n} - B_{T_{n-1},T_n-g_n}| \leq \frac{2M}{C_{n-1}}$ (D)

Finally, by periodicity;

$B_{T_{n-1},T_n} = \frac{1}{c_{n-1}}(f(\alpha_{n-1}) + \ldots + f(\sigma^{c_{n-1}-1}\alpha_{n-1})) = t_n$ (E)

Now, combining the estimates $(A), (B), (C), (D), (E)$, we have;

$|A_{T_n} - t_n| \leq \frac{M}{2^{n-2}} + \frac{2M}{2^{n-2}} + Q_{n-1} + \frac{2M}{C_{n-1}} = b_n$

Clearly $\{b_n\}_{n\in\mathcal{N}}$ is decreasing. Moreover, $lim_{n\rightarrow\infty}b_n = 0$, as $lim_{n\rightarrow\infty}C_n = \infty$, (iii), and by the choice of $\{Q_n\}_{n\in\mathcal{N}}$. This shows $(***)$. We now have to estimate the averages up to place between the critical points T_n and T_{n+1}.

Case 1. The place v is a periodic point of the form;

$T_n + mg_n$, where $0 \leq m \leq C_n - 1$

We have $A_v = \lambda A_{T_n} + (1-\lambda)A_{T_n,v}$ $(0 \leq \lambda \leq 1)$, where $|A_{T_n,v} - t_{n+1}| \leq Q_n$, by $(C), (E)$, and $|A_{T_n} - t_n| \leq b_n$, by $(***)$. Now, let $t = lim_{n\rightarrow\infty}t_n$. Given $\epsilon > 0$, choose $N(\epsilon)$, such that $|t_n - t| < \epsilon$, for all $n \geq N(\epsilon)$. Then;

$|A_v - t| \leq max\{|A_{T_n} - t|, |A_{T_n,v} - t|\}$

$\leq max\{b_n + \frac{\epsilon}{2}, Q_n + \frac{\epsilon}{2}\}$

Choose $N_1(\epsilon) \geq N(\epsilon)$, such that $max\{b_n, Q_n\} < \frac{\epsilon}{2}$, for all $n \geq N_1(\epsilon)$, then $|A_v - t| < \epsilon$, for all $n \geq N_1(\epsilon)$.

Case 2. The place v is a possibly non-periodic point of the form;

$T_n + w$, where $0 \leq w \leq T_{n+1} - T_{n-1} - g_n$.

Choose periodic points v_1 and v_2, with $T_n \leq v_1 \leq v \leq v_2 \leq T_{n+1} - g_n$, and $v_2 - v_1 = c_n$, so $0 \leq v - v_1 = e \leq c_n$. Then $A_v = \frac{v_1}{v_1+e}A_{v_1} + \frac{e}{v_1+e}A_{v_1,v}$. As $v_1 \geq T_n$, we have;

$\frac{e}{v_1+e} \leq \frac{e}{T_n+e} \leq \frac{c_n}{T_n} \leq \frac{1}{2^n}$ by (iv).

Therefore;

$|A_v - A_{v_1}| = |(1-\delta)A_{v_1} + \delta A_{v_1,v} - A_{v_1}|$, $(\delta \leq \frac{1}{2^n})$

$\leq \delta(|A_{v_1}| + |A_{v_1,v}|) \leq \frac{M}{2^{n-1}}$

For $n \geq N_1(\frac{\epsilon}{2})$, $|A_{v_1}-t| < \frac{\epsilon}{2}$, by Case 1, so $|A_v-t| < \epsilon$, for $n \geq N_2(\epsilon)$, where $N_2(\epsilon) = max\{N_1(\frac{\epsilon}{2}), log(\frac{2M}{\epsilon}) + 2\}$.

Case 3. The place v is of the form;

$T_n + w$, where $T_{n+1} - T_n - g_n \leq w \leq T_{n+1} - T_n$.

We have;

$A_v = \lambda A_{T_n} + (1-\lambda)A_{T_n,v}$, $(0 \leq \lambda \leq 1)$, $(\dagger)$,

$A_{T_n,T_{n+1}} = \mu A_{T_n,v} + (1-\mu)A_{v,T_{n+1}}$, $\frac{C_n-1}{C_n} \leq \mu \leq 1$

Therefore;

$|A_{T_n,T_{n+1}} - A_{T_n,v}| \leq \frac{2M}{C_n}$

$|A_{T_n,T_{n+1}} - t_{n+1}| \leq b_{n+1}$, by $(B),(C),(D),(E)$

$|A_{T_n,v} - t_{n+1}| \leq \frac{2M}{C_n} + b_{n+1}$

$|A_{T_n} - t_n| \leq b_n$ by $(***)$

$|A_v - t| \leq max\{|A_{T_n} - t|, |A_{T_n,v} - t|\}$ by $(\dagger)$

$\leq max\{b_n + |t_n - t|, \frac{2M}{C_n} + b_{n+1} + |t_{n+1} - t|\}$, $(\dagger\dagger)$

We have, for $n \geq N(\frac{\epsilon}{2})$, $max\{|t_n - t|, |t_{n+1} - t|\} < \frac{\epsilon}{2}$. Choose $N_3(\epsilon)$, such that $max\{b_n, \frac{2M}{C_n} + b_{n+1}\} < \frac{\epsilon}{2}$, for all $n \geq N_3(\epsilon)$. Then, for $n \geq N_3(\epsilon)$, $|A_v - t| < \epsilon$.

To complete the proof, let $N_4(\epsilon) = max\{N_1(\epsilon), N_2(\epsilon), N_3(\epsilon)\}$. Then, for $n \geq N_4(\epsilon)$, $|A_m - t| < \epsilon$, for all $m \geq T_n$, by Cases 1,2 and 3. Therefore;

$lim_{m\rightarrow\infty}\frac{1}{m}\sum_{i=0}^{m-1} f(\sigma^i\alpha) = \int_X f d\rho$

so α is typical, as required.

□

We now formulate the following criteria.

Lemma 6.12. *Suppose that for every $g \in C([0,1]^{\mathcal{N}})$, and $\epsilon > 0$, there exists a periodic element $\beta \in [0,1]^{\mathcal{N}}$, with;*

$$|\int_{[0,1]^{\mathcal{N}}} g d\rho_\beta - \int_{[0,1]^{\mathcal{N}}} g d\rho| < \epsilon$$

then there exists a sequence of periodic elements $(\alpha_n)_{n\in\mathcal{N}}$, with $(\rho_{\alpha_n})_{n\in\mathcal{N}}$ converging weakly to ρ.

Proof. We abbreviate $[0,1]^{\mathcal{N}}$ to X. Let $\mathcal{M}$ denote the vector space of real valued regular measures on $(X, \mathfrak{C})$. As we observed every probability measure belongs to $\mathcal{M}$. $\mathcal{M}$ is a Banach space, with norm defined by total variation, see [37]. Using the Riesz Representation Theorem, $\mathcal{M}$ can be identified with the dual space $C(X)^*$. It is easy to see that then $\mathcal{M} \cong C(X)^*$, as Banach spaces, however, we will not require this fact. The weak $*$-topology, see [9], on $\mathcal{M}$, is the coursest topology for which all the elements $\hat{g} \in C(X)^{**}$, where $g \in C(X)$, are continuous. Formally, we define a set $U \subset \mathcal{M}$ to be open if for all $\rho \in U$, there exist $\{g_1, \ldots, g_n\} \subset C(X)$, and positive reals $\{\epsilon_1, \ldots, \epsilon_n\}$ such that;

$$\{\rho' \in \mathcal{M} : |\rho'(g_i) - \rho(g_i)| < \epsilon_i\} \subset U$$

Fixing ρ, let Ω_ρ denote the open sets containing ρ. We show that Ω_ρ has a countable base, $(*)$. Using the compactness argument, given in Lemma 6.11, and the Stone-Weierstrass Theorem, see [9], it is easy to show that the space V of pullbacks of polynomial functions on $[0,1]^n$, for some n, is dense in $C(X)$. Clearly V has a countable basis, which shows that $C(X)$ is separable, that is, contains a countable dense subset Y. Now suppose that $g \in C(X)$, $\epsilon > 0$. Let $U_{g,\epsilon} = \{\rho' : |\rho'(g) - \rho(g)| < \epsilon\}$, and $D \in \mathcal{Q}$. Choose $\delta \in \mathcal{Q}$ with $\delta < \frac{\epsilon}{2(D+2|\rho(X)|)}$, and $\gamma \in \mathcal{Q}$ with $\gamma < \frac{\epsilon}{2}$. Choose $h \in Y$ with $||g-h||_{C(X)} < \delta$. Then $U_{h,\gamma} \cap U_{1,D} \subset U_{g,\epsilon}$, $(**)$, as if $|\rho'(h) - \rho(h)| < \gamma$, then;

$$|\rho'(g) - \rho(g)| = |\rho'(g-h) + \rho'(h) - \rho(g-h) - \rho(h)| \leq \delta(|\rho'(X)| + |\rho(X)|) + \gamma$$

and, if $|\rho'(1) - \rho(1)| < D$, then $|\rho'(X)| + |\rho(X)| < D + 2|\rho(X)|$, so $|\rho'(g) - \rho(g)| < \epsilon$. This clearly shows $(**)$. As sets of the form $U_{h,q} \in \Omega_\rho$, for $h \in Y$, and $q \in \mathcal{Q}$, are countable, we clearly have $(*)$. Let $I : \mathcal{N} \to \Omega_\rho$ be an enumeration of the sets $U_{h,q}$, and let $J : \mathcal{N} \to \Omega_\rho$ define the intersection of the first n elements in I. If the assumption in the lemma is satisfied, we can define a sequence of probability measures

$(\rho_{\alpha_n})_{n\in\mathcal{N}}$, by taking ρ_{α_n} to lie inside the open set $J(n)$. Then clearly such a sequence converges to ρ in the weak $*$-topology, hence, for any $g \in C(X)$, as g is continuous for this topology $lim_{n\rightarrow\infty}\rho_{\alpha_n}(g) = \rho(g)$. Therefore, the sequence $(\rho_{\alpha_n})_{n\in\mathcal{N}}$ converges weakly to ρ. □

We refine this criteria further;

Definition 6.13. *Given a positive integer m, we define the partition E_m of $[0,1]$ to consist of the sets;*

$E_{j,m} = [\frac{j}{m}, \frac{j+1}{m})$ *for j an integer between 0 and $m-2$*

$E_{m-1,m} = [\frac{m-1}{m}, 1]$

Given positive integers m, n, we define the partition $B_{m,n}$ of $[0,1]^n$ to consist of the sets;

$B_{\bar{j},m,n} = E_{j_0,m} \times E_{j_1,m} \times \ldots \times E_{j_{n-1},m}$

where $\bar{j} = (j_0, j_1, \ldots, j_{n-1})$ and $\{j_0, \ldots, j_{n-1}\}$ are integers between 0 and $m-1$.

We define the partition $C_{m,n}$ of $[0,1]^{\mathcal{N}}$ to consist of the sets;

$C_{\bar{j},m,n} = \pi_n^{-1}(B_{\bar{j},m,n})$

where π_n is the projection onto the first n coordinates.

Lemma 6.14. *Let $\epsilon > 0$, $g \in C(X)$ be given as in Lemma 6.12, and let ρ' be a regular Borel measure, then there exist positive integers m, n, and $\delta > 0$, such that, if;*

$|\rho'(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| < \delta$

for all sets $C_{\bar{j},m,n}$ belonging to $C_{m,n}$, then;

$|\int_{[0,1]^{\mathcal{N}}} g d\rho' - \int_{[0,1]^{\mathcal{N}}} g d\rho| < \epsilon$

Proof. For a positive integer n, let W_n consist of the inverse images in X (from the projection π_n) of open boxes in $[0,1]^n$, with rational corners. Let $W = \bigcup_{n\in\mathcal{N}} W_n$. It is clear that W forms a countable basis for the topology on $[0,1]^{\mathcal{N}}$. Adapting the compactness argument, given above in Lemma 6.11, for any $\gamma > 0$ and $g \in C(X)$, we can find a positive integer n, and finitely many sets $\{W_{1,n}, \ldots, W_{r,n}\}$ in W_n, covering X, such that $|g(x) - g(y)| < \gamma$ for all x, y in $W_{j,n}$, $1 \leq j \leq r$. Now choose m such that each set of the partition $C_{m,n}$ lies inside one of the $W_{j,n}$. Then $|g(x) - g(y)| < \gamma$ on each $C_{\bar{j},m,n}$, belonging to $C_{m,n}$. Now, for given $\delta > 0$, suppose we choose ρ' such that $|\rho'(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| < \delta$, $(*)$. Then;

$$|\int_X g d\rho' - \int_X g d\rho| = |\sum_{\bar{j}} \int_{C_{\bar{j},m,n}} g d\rho' - \sum_{\bar{j}} \int_{C_{\bar{j},m,n}} g d\rho|$$

$$\leq \sum_{\bar{j}} |\int_{C_{\bar{j},m,n}} g d\rho' - \int_{C_{\bar{j},m,n}} g d\rho|, \, (**)$$

Without loss of generality, assuming ρ' is positive, by definition of the integral, see [37], we have that;

$$c_{\bar{j}} \rho'(C_{\bar{j},m,n}) \leq \int_{C_{\bar{j},m,n}} g d\rho' \leq d_{\bar{j}} \rho'(C_{\bar{j},m,n})$$

$$c_{\bar{j}} \rho(C_{\bar{j},m,n}) \leq \int_{C_{\bar{j},m,n}} g d\rho \leq d_{\bar{j}} \rho(C_{\bar{j},m,n})$$

where $c_{\bar{j}} = inf_{C_{\bar{j},m,n}} g$ and $d_{\bar{j}} = sup_{C_{\bar{j},m,n}} g$. Then;

$$c_{\bar{j}} \rho'(C_{\bar{j},m,n}) - d_{\bar{j}} \rho(C_{\bar{j},m,n}) \leq \int_{C_{\bar{j},m,n}} g d\rho' - \int_{C_{\bar{j},m,n}} g d\rho$$

$$\leq d_{\bar{j}} \rho'(C_{\bar{j},m,n}) - c_{\bar{j}} \rho(C_{\bar{j},m,n})$$

Therefore, again, without loss of generality;

$$|\int_{C_{\bar{j},m,n}} g d\rho' - \int_{C_{\bar{j},m,n}} g d\rho|$$

$$\leq (d_{\bar{j}} - c_{\bar{j}}) \rho'(C_{\bar{j},m,n}) + |c_{\bar{j}}||\rho'(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| \leq \gamma \rho'(C_{\bar{j},m,n}) + |c_{\bar{j}}|\delta$$

$(***)$

By $(*)$, $\rho'(X) = \sum_{\bar{j}} \rho'(C_{\bar{j},m,n}) \leq \sum_{\bar{j}} \rho(C_{\bar{j},m,n}) + \delta m^n = 1 + \delta m^n$, so using $(**)$, $(***)$, and the fact that $|g| \leq M$;

$$|\int_X g d\rho' - \int_X g d\rho| \leq \gamma(1 + \delta m^n) + \delta M m^n$$

So if we choose $0 < \gamma < \frac{\epsilon}{2}$ and $0 < \delta < \frac{\epsilon}{2(\gamma + M) m^n}$, we obtain;

$|\int_X g d\rho' - \int_X g d\rho| < \epsilon$

as required.

□

We finally claim;

Theorem 6.15. *If $C_{m,n}$ is a partition, as in Definition 6.13 and $\delta > 0$, then there exists a periodic element β, such that;*

$|\rho_\beta(C_{\bar{j},m,n}) - \rho(C_{\bar{j},m,n})| < \delta$

for all sets $C_{\bar{j},m,n}$ belonging to $C_{m,n}$.

Proof. Let $\Sigma = \{\frac{1}{2m}, \frac{3}{2m}, \ldots, \frac{2m-1}{2m}\}$. Define $\kappa : \Sigma^n \rightarrow \mathcal{R}$ by;

$\kappa((\frac{2j_0+1}{2m}, \ldots, \frac{2j_{n-1}+1}{2m})) = \rho(C_{\bar{j},m,n})$

As $C_{m,n}$ is a partition of X and ρ is a probability measure, κ is a probability measure on Σ^n. Moreover, using the partition property and the fact that ρ is σ-invariant;

$\sum_{\xi_0 \in \Sigma} \kappa((\xi_0, \ldots, \xi_{n-1})) = \rho(\pi_n^{-1}([0,1] \times E_{j_1,m} \times \ldots \times E_{j_{n-1},m}))$

$= \rho(\pi_n^{-1}(E_{j_1,m} \times \ldots \times E_{j_{n-1},m} \times [0,1]))$

$= \sum_{\xi_0 \in \Sigma} \kappa((\xi_1, \ldots, \xi_{n-1}, \xi_0))$ $(*)$

Now let $N > 0$ be a sufficiently large positive integer, then we claim that we can find a probability measure κ' on Σ^n such that;

(i). $|\kappa'(\bar{\xi}) - \kappa(\bar{\xi})| < \delta$

(ii). The condition $(*)$ still holds.

(iii). $N\kappa'(\bar{\xi})$ is a non-negative integer, for all $\bar{\xi} \in \Sigma^n$

This follows from a simple linear algebra argument. We can identify the set of real measures on Σ^n with the real vector space V of dimension m^n. The condition $(*)$ then defines a subspace $W \subset V$. The condition of being a probability measure requires that;

$\sum_{\xi_0,\ldots,\xi_{n-1}\in\Sigma^n} \kappa((\xi_1,\ldots,\xi_{n-1},\xi_0)) = 1$, $(**)$

which defines an affine space $S_{aff} \subset V$. $S_{aff} \cap W$ contains a rational point q, corresponding to the probability measure with coordinates m^{-n}. It is straightforward to see that $(S_{aff} \cap W) = [(S_{aff} - q) \cap W] + q$. Moreover, $(S_{aff} - q) \cap W$ is a vector space defined by rational coefficients, so it has a rational basis. This shows that rational points are dense in $S_{aff} \cap W$. We can, without loss of generality, assume that all the coordinates of κ are strictly greater than zero. If not, consider instead the space $S_{aff} \cap W \cap W'$, where $W' = Ker(\pi)$ is the kernel of the projection onto the non-zero coordinates of κ. The same argument shows that rational points are dense in $S_{aff} \cap W \cap W'$. We can now obtain a probability measure κ', satisfying conditions $(i) - (iii)$, by finding a rational vector sufficiently close to κ in $S_{aff} \cap W$, and choosing N large enough.

Now take a longest sequence $\{\xi^0, \ldots, \xi^{r-1}\}$ of elements in Σ^n, such that;

(1). $(\xi_1^i, \ldots, \xi_{n-1}^i) = (\xi_0^{i+1}, \ldots, \xi_{n-2}^i)$.

(2). $Card(\{i : 0 \leq i < r, \xi^i = \xi\}) \leq N\kappa'(\xi)$ for any $\xi \in \Sigma^n$

where $\xi^i = (\xi_0^i, \ldots, \xi_{n-1}^i)$, for $0 \leq i \leq r$, and $\xi^r = \xi^0$.

Then, by graph theoretical considerations, ([22]), one can show that equality holds in the above inequality in (2), for any $\xi \in \Sigma^n$, $(* * *)$.

[22]The graph theory argument proceeds as follows. We construct a tree. For every $\xi' \in \Sigma^{n-1}$, where $\xi' = (\xi_1, \ldots, \xi_{n-1})$, associate a vertex $v_{\xi'}$ (the trunk). Similarly, for every $\xi \in \Sigma^n$, where $xi = (\xi_0, \ldots, \xi_{n-1})$, associate two vertices l_ξ (left) and r_ξ (right). Attach the vertex l_ξ to $v_{\xi'}$ iff $\pi(\xi) = \xi'$, where π is the projection onto the last $n-1$ coordinates, and, attach l_ξ to $v_{\xi'}$. iff $\pi'(\xi) = \xi'$, where π' is the projection onto the first $n-1$ coordinates. In this way, we obtain a tree, having $m^{n-1}(2m+1)$ vertices, $m^{n-1}(2m)$ branches, and m_{n-1} components. Each element $\xi \in \Sigma^n$ corresponds to two vertices, one on the left and one on the right of the tree. Now attach weights $m_\xi = n_\xi$ to the left vertices and right vertices respectively, by assigning the vertices l_ξ and r_ξ, the weights $m_\xi = N\kappa'(\xi)$ and $n_\xi = N\kappa'(\xi)$ respectively. Observe that, by the condition $(*)$ in the main text, for any given ξ';

$m_{\xi'} = \sum_{\xi \in \Sigma^n : \pi(\xi) = \xi'} m_\xi = n_{\xi'} = \sum_{\xi \in \Sigma^n : \pi'(\xi) = \xi'} n_\xi$ $(\dagger)$

Now, given a sequence $\{\xi^0, \xi^1, \ldots, \xi^k\}$ of elements in Σ^n, where $\xi^i = (\xi^i_0, \ldots, \xi^i_{n-1})$, for $0 \leq i \leq k$, we attach sets L_ξ to each vertex l_ξ, by requiring that, $\xi^i \in L_\xi$ iff $\xi^i = \xi$, and, similarly, we attach sets R_ξ to each vertex r_ξ. We call a sequence allowed if (i). For each $\xi \in \Sigma^n$, $Card(L_\xi) = Card(R_\xi) \leq m_\xi = n_\xi$ and (ii). For each $1 \leq i \leq k$, if ξ^i appears in the set R_ξ, then ξ^{i-1} appears in a set $L_{\xi''}$, where $l_{\xi''}$ and r_ξ are attached to the same vertex $v_{\xi'}$, so that $\pi(\xi'') = \pi'(\xi) = \xi'$. Clearly, all allowed sequences are bounded in length by $N\kappa'(X)$, so there exists a longest allowed sequence $s = (\xi^i)_{0 \leq i \leq t}$. Let ξ^t be the final element in the sequence, and suppose that $\xi^t \in L_{\xi''}$, then, we claim that ξ^0 belongs to a set R_ξ, where $\pi(\xi'') = \pi'(\xi) = \xi'$, $(\dagger\dagger)$. If not, all such sets R_ξ, with $\pi'(\xi) = \pi(\xi'')$, consists of elements ξ^i with $i \geq 1$. If, for one of these sets R_ξ, $Card(R_\xi) \lneq n_\xi$, then we can extend the sequence by setting $\xi^{t+1} = \xi$, clearly such a sequence is allowed, contradicting maximality. So we can assume that $Card(R_\xi) = n_\xi$. By condition (ii), for every element ξ^i, $i \geq 1$, appearing in R_ξ, there exists an element ξ^{i-1} appearing in an $L_{\xi''}$, with $\pi(\xi'') = \pi(\xi^t)$. This provides a total of $w+1$ elements appearing in such $L_{\xi''}$, where $w = \sum_{\xi \in \Sigma^n : \pi'(\xi) = \xi'} n_\xi$. By $(\dagger)$, this is greater than $\sum_{\xi \in \Sigma^n : \pi(\xi) = \xi'} m_\xi$. Clearly, this contradicts condition (i) of an allowed path. Hence, $(\dagger\dagger)$ is shown. Observe also that if $\xi' \in \Sigma^{n-1}$, and $s_{r,\xi'}$ denotes the total number of elements from the sequence s, appearing in sets to the right of ξ', $s_{l,\xi'}$, to the left, then $s_{l,\xi'} = s_{r,\xi'}$, In particular, by $(\dagger)$, $m_{\xi'} - s_{l,\xi'} = n_{\xi'} - s_{r,\xi'} \geq 0$, so the number of "vacant slots" (if there are any), is the same on both sides of a given ξ', $(\dagger\dagger\dagger)$. In order to see this, we can, without loss of generality, assume that $\pi'(\xi^0) \neq \xi'$, then just note that an element ξ^{i+1} belongs to a set on the right of ξ' iff ξ^i belongs to a set on the left of ξ', by condition (ii) of an allowed path. We now claim that for all $\xi \in \Sigma^n$, $Card(R_\xi) = n_\xi$, $(\dagger\dagger\dagger\dagger)$, (so there are no vacant slots). We have already shown this in the particular case when $\pi'(\xi) = \pi'(\xi^0)$. We define an element ξ to be cyclic if $\pi(\xi) = \pi'(\xi)$, so cyclic elements are just constant sequences. We define an element ξ to be free if $Card(R_\xi) \lneq n_\xi$. No free cyclic element ξ_{cyc} can encounter the sequence s, for suppose that there exists a ξ^i, for some $0 \leq i \leq t$, with $\pi(\xi^i) = \pi'(\xi_{cyc})$, then we can extend the sequence s to $s' = \{\xi^0, \ldots, \xi^i, \xi_{cyc}, \xi^{i+1}, \ldots, \xi^t\}$, and still obtain an allowed path, contradicting maximality. So we have that, if ξ is free cyclic, with $\pi_\xi = \xi'$, then $s_{l,\xi'} = s_{r,\xi'} = 0$, $(\dagger\dagger\dagger\dagger\dagger)$. Now suppose there exists a

Now let β be the periodic element in $[0,1]^{\mathcal{N}}$, with period $n+r-1$, defined by;

$(\beta(0),\beta(1),\ldots,\beta(n+r-2))=(\xi_0^0,\xi_1^0,\ldots,\xi_{n-1}^0,\xi_{n-1}^1,\xi_{n-1}^2,\ldots,\xi_{n-1}^{r-1})$

By (i), it is sufficient to prove that, for each $\bar{j}\in m^n$;

$|\rho_\beta(C_{\bar{j},m,n})-\kappa'(\xi_{\bar{j}})|<\epsilon$, $(****)$,

where $\epsilon=min_{\bar{i}}(\delta-|\kappa'(\xi_{\bar{i}})-\kappa(\xi_{\bar{i}})|)$, and $\xi_{\bar{j}}$ is the unique element of Σ^n lying inside $C_{\bar{j},m,n}$. By definition of ρ_β, $\rho_\beta(C_{\bar{j},m,n})=\frac{c_{\bar{j}}}{n+r-1}$, where;

$c_{\bar{j}}=Card(\{k:0\leq k<n-r-1,\pi_n(\sigma^k(\beta))=\xi_{\bar{j}}\})$.

By definition of β, and $(***)$, $c_{\bar{j}}=\frac{N\kappa'(\xi_{\bar{j}})+y}{n+r-1}$, where $0\leq y\leq n$. As κ' is a probability measure, again by $(***)$, we have that $r-1=N$. Hence;

$\frac{c_{\bar{j}}}{n+r-1}=\frac{N\kappa'(\xi_{\bar{j}})+y}{N+n}=\kappa'(\xi_{\bar{j}})+\frac{y-n\kappa'(\xi_{\bar{j}})}{N+n}$.

Therefore,

$|\rho_\beta(C_{\bar{j},m,n})-\kappa'(\xi_{\bar{j}})|\leq\frac{n}{N+n}<\epsilon$.

---

free element ξ_{free}. Choose the largest k, with $0\leq k\leq t$, such that ξ^k appears in $L_{\xi''}$ with $\pi(\xi'')=\pi'(\xi_{free})$, $(\sharp)$. As we have observed, $k\lneq t$. We construct a forward path from ξ_{free} as follows. Define $\eta^0=\xi_{free}$, add the element η^0 to $R_{\xi_{free}}$ and $L_{\xi_{free}}$, and call the new sets $R_{0,\xi}$ and $L_{0,\xi}$, for $\xi\in\Sigma^n$. Having defined η^j, there are four cases. If $\pi(\eta^j)=\pi'(\eta^0)$, terminate the sequence. Otherwise, if $\pi(\eta^j)=\pi(\xi_{cyc})$ for some cyclic element with $Card(R_{j,\xi_{cyc}})\lneq n_{\xi_{cyc}}$, then define $\eta^{j+1}=\xi_{cyc}$, add the element η^{j+1} to $R_{j,\xi_{cyc}}$ and $L_{i,\xi_{cyc}}$, calling the new sets $R_{j+1,\xi}$ and $L_{j+1,\xi}$, for $\xi\in\Sigma^n$. If there is no such cyclic element, and there exists a free element ξ' with $\pi(\eta^j)=\pi'(\xi')$ and $Card(R_{j,\xi'})\lneq n_{\xi'}$, then define $\eta^{j+1}=\xi'$ (so there is some choice here), and, as before, redefine the sets $R_{j,\xi}$ and $L_{j,\xi}$ to $R_{j+1,\xi}$ and $L_{j+1,\xi}$, for $\xi\in\Sigma^n$. If there is no free element of this form, then terminate the sequence. It is straightforward to see, using $(\dagger\dagger\dagger)$, $(\dagger\dagger\dagger\dagger\dagger)$, and the fact that η^0 is not cyclic, that the sequence $\{\eta^0,\ldots,\eta^j\}$ terminates after a finite number of steps l, with $l>0$, and $\pi(\eta^l)=\pi'(\eta^0)$. Moreover, for all $k<i<t$, and $0\leq j\leq l$, we have that $\pi(\xi^i)\neq\pi'(\eta^j)$, by $(\sharp)$. Hence, we can construct an allowed sequence $s''=\{\xi^0,\ldots,\xi^k,\eta^0,\ldots,\eta^l\xi^{k+1},\ldots,\xi^t\}$, contradicting maximality of s. This shows $(\dagger\dagger\dagger\dagger)$. It is clear that the sequence $s'''=\{\xi^0,\ldots,\xi^{r-1}\}$, as defined in the main text, is a longest allowed sequence, as defined in this footnote, using $(\dagger\dagger)$. Hence, by $(\dagger\dagger\dagger\dagger)$, we have equality in (2) as required.

if we choose N sufficiently large. Hence, $(****)$ and the theorem are shown.

□

We summarise what we have done;

Theorem 6.16. *The Ergodic Theorem (6.1) holds and admits a nonstandard proof.*

Proof. Combine Theorems 6.3,6.9,6.15, and Lemmas 6.4,6.6,6.7,6.11,6.12,6.14.

□

Remarks 6.17. *There are some outstanding questions in Ergodic Theory, which one might hope to solve using nonstandard methods, similar to the above. One of these is Ornstein's Isomorphism Theorem, I hope to investigate this direction further.*

7. Stochastic Calculus

We use the same notation as in Section 3. $(X, \mathfrak{A}, \nu)$ is a *hyperfinite* probability space, ([23]), and we let $(X, \mathfrak{M}_L, \mu_L)$ be the corresponding Loeb space, see Lemma 3.15.

We make the following definition;

Definition 7.1. *We define a random variable on $(X, \mathfrak{A}, \nu)$ to be a function $x : X \rightarrow {}^*\mathcal{R}$, which is $\mathfrak{A}$-measurable.*

An internal collection of random variables $\{x_i\}_{i \in I}$ is $*$-independent if, for every $*$-finite internal subcollection $\{x_1, \ldots, x_m\}$ and every internal m-tuple $(\alpha_1, \ldots, \alpha_m) \in {}^*\mathcal{R}^m$, with $m \in {}^*\mathcal{N}$;

$$\nu(\{\omega : x_1(\omega) < \alpha_1, \ldots, x_m(\omega) < \alpha_m\}) = \prod_{k=1}^{m} \nu(\{\omega : x_k(\omega) < \alpha_k\})$$

An internal collection of random variables $\{x_i\}_{i \in I}$ is S-independent if, for every finite subcollection $\{x_1, \ldots, x_m\}$ and every m-tuple $(\alpha_1, \ldots, \alpha_m) \in \mathcal{R}^m$, with $m \in \mathcal{N}$;

$$\nu(\{\omega : x_1(\omega) < \alpha_1, \ldots, x_m(\omega) < \alpha_m\}) \simeq \prod_{k=1}^{m} \nu(\{\omega : x_k(\omega) < \alpha_k\})$$

[23] For the results, not explicitly involving Brownian motion, we can sometimes (not the expectation arg below, but CLT OK) weaken this assumption to $*$-sigma algebra

Observe that $*$-independence implies S-independence, ([24]).

We define two random variables to be strongly equivalent if;

$\nu(\{\omega : x_1(\omega) = x_2(\omega)\}) = 1.$

Let $\mathfrak{A}'$ be an internal subalgebra of $\mathfrak{A}$, and x a random variable, then we define the conditional expectation $E(x|\mathfrak{A}')$ to be the strong equivalence class of $\mathfrak{A}'$-measurable random variables y, for which;

$\int_{A'} y d\nu = \int_{A'} x\nu$, *for all $A' \in \mathfrak{A}'$, as in Definition 3.9, $(*)$.*

We define $E(x) = E(x|\mathfrak{A}')$, where $\mathfrak{A}'$ is the internal subalgebra $\{\emptyset, X\}$, ([25]).

Lemma 7.2. *Suppose that $\{x_i\}_{i\in I}$ is an S-independent collection of random variables on $(X, \mathfrak{A}, \nu)$, then $\{{}^\circ x_i\}_{i\in I}$ is an independent, ([26]), collection of random variables on $(X, \mathfrak{M}_L, \mu_L)$.*

Proof. Suppose that $m \in \mathcal{N}_{>0}$ and $(\alpha_1, \ldots, \alpha_m) \in \mathcal{R}^m$, then;

$\mu_L(\{\omega : {}^\circ x_1(\omega) < \alpha_1, \ldots, x_m(\omega) < \alpha_m\})$

$= lim_{n\rightarrow\infty} {}^\circ\nu(\{\omega : x_1(\omega) < \alpha_1 - \frac{1}{n}, \ldots, x_m(\omega) < \alpha_m - \frac{1}{n}\})$

$= lim_{n\rightarrow\infty} {}^\circ(\prod_{k=1}^m \nu(\{\omega : x_k(\omega) < \alpha_k - \frac{1}{n}\}))$

[24] All the proofs go through if you weaken the condition of $*$-independence to $\simeq$, but this is how it is defined in the paper [1].

[25] The facts that there exists a $\mathfrak{A}'$-measurable random variable y satisfying the property $(*)$, and, it is unique up to strong equivalence, follow immediately by transfer from the corresponding fact for finite algebras, see [44]. If two random variables x_1 and x_2 are strongly equivalent, then, again by transfer, $E(x_1|\mathfrak{A}') = E(x_2|\mathfrak{A}')$. One could also define the two random variables to be weakly equivalent if $\nu(\{\omega : x_1(\omega) = x_2(\omega)\}) \simeq 1$. However, it is not true then that $E(x_1|\mathfrak{A}') = E(x_2|\mathfrak{A}')$. This can be seen by taking X to be a $*$- finite interval $[0, \eta]$, with counting measure ν on the $*$-algebra $\mathfrak{A}$ of all internal subsets. If $x_1 = 0$ and $x_2 = 0$, for $0 < n < \eta$, $x_2(\eta) = 1$, then $E(x_1) = 0$, but $E(x_2) = \frac{1}{\eta}$. Note that 0 and $\frac{1}{\eta}$ are not even weakly equivalent. We always take the standard definition of conditional expectation to mean an equivalence class, *a.e.* Of course, all the standard properties of conditional expectations, see [44] for a comprehensive list, transfer to the hyperfinite case.

[26] See [44] for a standard definition of independence. Observe that a definition involving countable infinite products (in the standard sense) is equivalent to the finite product definition. We *don't* need $*$-independence to show this.

$= \prod_{k=1}^{m} lim_{n\rightarrow\infty} {}^{\circ}\nu(\{\omega : x_k(\omega) < \alpha_k - \frac{1}{n}\})$

$= \prod_{k=1}^{m} \mu_L(\{\omega : {}^{\circ}x_k(\omega) < \alpha_k\})$

The first and last lines follows easily from the facts that ${}^{\circ}x_k(\omega) < \alpha_k$ iff $x_k(\omega) < \alpha_k - \frac{1}{n}$ for some $n \in \mathcal{N}$, and μ_L defines a countably additive measure. The second line is just the definition of S-independence. The rest just uses elementary properties of the standard part mapping and limits.

□

Theorem 7.3. *Let $\mathfrak{A}'$ be an internal subalgebra of $\mathfrak{A}$, then there exists a unique internal partition P of X, generating $\mathfrak{A}'$. Let $(X, \mathfrak{M}'_L, \mu'_L)$ denote the Loeb space corresponding to $(X, \mathfrak{A}', \nu)$. Then;*

(i). If g is $\mathfrak{M}_L$-measurable and $g|A$ is constant for each $A \in P$, then g is $\mathfrak{M}'_L$-measurable.

(ii). If $f \in SL^p(X, \mathfrak{A}, \nu)$, then $E(f|\mathfrak{A}') \in SL^p(X, \mathfrak{A}', \nu)$ and;

${}^{\circ}E(f|\mathfrak{A}') = E({}^{\circ}f|\mathfrak{M}'_L)$, ([27]).

.

Proof. The fact that there exists a unique internal partition P of X, generating $\mathfrak{A}'$, follows immediately from the corresponding fact for finite algebras.

(i). Let $\alpha \in \mathcal{R}$ and let $B = \{x : g(x) < \alpha\}$. Let $\epsilon > 0$ be standard, then By Lemma 3.15 (3.4)(i,ii), we can find C and D in $\mathfrak{A}$, with $C \subset B \subset D$, such that $\nu(D \setminus C) < \epsilon$. For $x \in X$, let A_x be the unique element of P such that $x \in A_x$. Let $D' = \{x \in X : A_x \subset D\}$, and, let $C' = \{x \in X : A_x \cap C \neq \emptyset\}$. Then $C \subset C' \subset B \subset D' \subset D$, as $g|A_x$ is constant, for $x \in X$. Moreover C' and D' belong to $\mathfrak{A}'$, they are just $*$-finite unions of sets in P. Therefore $\nu(D' \setminus C') < \epsilon$. As ϵ was arbitrary, $B \in \mathfrak{M}'_L$. This follows from the fact that $(X, \mathfrak{M}'_L, \mu'_L)$ is complete, and noting that $C'' = \bigcup_{i\in\mathcal{N}} C'_i$, and $D'' = \bigcap_{i\in\mathcal{N}} D'_i$ belong to $\mathfrak{M}'$, where $C'_i \subset B \subset D'_i$ are chosen such that $\nu(D'_i \setminus C'_i) < \frac{1}{i}$, C'_i and D'_i belong to $\mathfrak{A}'$.

[27]Here, $E(f|\mathfrak{A}')$ is the equivalence class of functions in $SL^p(X, \mathfrak{A}', \nu)$, see Definition 3.32, of a representative of $E(f'|\mathfrak{A}')$, as in Definition 7.1, for a representative f' of the equivalence class f! That this is a good definition will become clear from the proof. The rest of the notation is clear from Theorem 3.34.

(ii). Let f' be a representative of f, and let g be a representative of $E(f'|\mathfrak{A}')$. If $A \in P$, then g is constant on A. This follows as g is $\mathfrak{A}'$-measurable, so if $x \in A$, $g^{-1}(g(x))$, belongs to $\mathfrak{A}'$. Therefore, $g^{-1}(g(x))$ can be written, by transfer, as a $*$-finite union of sets in P, clearly such a union must contain A_x. If $\nu(A) > 0$, then, using Definition 7.1, we have;

$$g(x)\nu(A) = \int_A g d\nu = \int_A f' d\nu$$

so that;

$$g(x) = \frac{1}{\nu(A)} \int_A f' d\nu \ (*)$$

Now, if $A' \in \mathfrak{A}'$, with $\nu(A') > 0$, write A' as a $*$-finite union $\bigcup_{n=1}^{\omega} A_n$, where $A_n \in P$ and $\nu(A_n) > 0$. Then, using the fact that P is a partition, $(*)$, and Holder's inequality, see [37];

$$\int_{A'} |g|^p d\nu = \sum_{n=1}^{\omega} \int_{A_n} |g|^p d\nu$$

$$= \sum_{n=1}^{\omega} \nu(A_n) \left|\frac{1}{\nu(A_n)} \int_{A_n} f' d\nu\right|^p \leq \sum_{n=1}^{\omega} \int_{A_n} |f'|^p d\nu = \int_{A'} |f'|^p d\nu \ (**)$$

This clearly implies that $[g] \in SL^p(X, \mathfrak{A}', \nu)$, as $f \in SL^p(X, \mathfrak{A}, \nu)$, and if $A'' \in \mathfrak{A}'$, with $\nu(A'') = 0$, then we always have that $\int_{A''} |g|^p d\nu = 0$. This also requires the fact that the above calculation holds, if we replace $\int_{A'} |g|^p d\nu$ by $\int_{A'} |g|^p d(\nu|_{\mathfrak{A}})$; the two integrals are equal, see Lemma 3.15.

We can also conclude that $[g] \in SL^p(X, \mathfrak{A}, \nu)$, as, if K is infinite, then the set $\{|g|^p > K\}$ belongs to $\mathfrak{A}'$, and $\nu(\{|g|^p > K\}) \simeq 0$, otherwise $\int_X |g|^p d\nu$ would not be finite. By $(**)$ and the fact that $f \in SL^p(X, \mathfrak{A}, \nu)$, $\int_{\{|g|^p > K\}} |g|^p d\nu \simeq 0$. Now use Lemma 3.19.

By Theorem 3.34(iv), $[g] \in SL^1(X, \mathfrak{A}', \nu)$, $[g] \in SL^1(X, \mathfrak{A}, \nu)$ and $[f'] \in SL^1(X, \mathfrak{A}, \nu)$. Now, if $B' \in \mathfrak{M}'_L$, we can, by Lemma 3.15(3.4(ii)), find $B \in \mathfrak{A}'$, such that $\mu'_L(B \Delta B') = \mu_L(B \Delta B') = 0$. Then, by Theorem 3.34(i) and the fact that $[g] = E(f'|\mathfrak{A}')$;

$$\int_{B'} {}^\circ g d\mu_L = \int_B {}^\circ g d\mu_L = {}^\circ \int_B g d\nu = {}^\circ \int_B f' d\nu = \int_B {}^\circ f' d\mu_L = \int_{B'} {}^\circ f' d\mu_L$$

As ${}^\circ g$ is $\mathfrak{M}'_L$-measurable, by Theorem 3.15(3.8), we conclude that $[{}^\circ g] = E({}^\circ f'|\mathfrak{M}'_L)$. Now, running back through the calculation, it is easy to see that, for a given f', if I had chosen another representative g' of $E(f'|\mathfrak{A}'$, then, as g and g' are strongly equivalent, I would still obtain the same class $[g'] \in SL^p(X, \mathfrak{A}', \nu)$ and $[{}^\circ g'] = E({}^\circ f'|\mathfrak{M}'_L)$. Moreover, if I had chosen another representative f'' of f, then by Theorem 3.34(i), $E({}^\circ f'|\mathfrak{M}'_L) = E({}^\circ f''|\mathfrak{M}'_L)$. Now, taking g corresponding to f', as in the proof, and g'' corresponding to f'', repeat the calculation with $g - g''$ replacing g and $f - f''$ replacing f. It is easy to then conclude that $[g]$ and $[g'']$ determine the same class in $SL^p(X, \mathfrak{A}', \nu)$, hence in $SL^1(X, \mathfrak{A}', \nu)$. Again, by Theorem 3.34(i), we obtain $[{}^\circ g] = [{}^\circ g''] = E({}^\circ f'|\mathfrak{M}'_L)$.

□

Theorem 7.4. *Non-Standard Central Limit Theorem*

Suppose $\{x_n\}_{n \in {}^*\mathcal{N}}$ is an internal sequence of $*$-independent random variables on $(X, \mathfrak{A}, \nu)$, with common standard *distribution function F, ([28]). Suppose further that $E(x_n) = 0$ and $E(x_n^2) = 1$. Then, for every $n \in {}^*\mathcal{N} \setminus \mathcal{N}$, and $\alpha \in {}^*\mathcal{R}$;*

$$\nu(\{\omega : \tfrac{1}{\sqrt{n}} \textstyle\sum_{k=1}^{n} x_k(\omega) \leq \alpha\}) \simeq {}^*\Psi(\alpha)$$

where Ψ is the standard Gaussian distribution.

Proof. Let G be the distribution function of one of the standard random variables ${}^\circ x_n$ on $(X, \mathfrak{M}_L, \mu_L)$. Fix $\alpha \in \mathcal{R}$ and $\epsilon > 0$ standard. Since F is $*$-right continuous, there exists $m \in {}^*\mathcal{N}$, with $F(\alpha + \frac{1}{r}) < F(\alpha) + \epsilon$, $(*)$ for all $r \geq m$. Since F is the transfer of a standard function, we can take $m \in \mathcal{N}$, such that $(*)$ holds for all $r \geq m$, with $r \in \mathcal{N}$. As ϵ was arbitrary, this means that $lim_{m \to \infty} F(\alpha + \frac{1}{m}) = F(\alpha)$, $(**)$. Then, using the definition of G, a similar calculation to Lemma 7.2, and $(**)$;

$$G(\alpha) = \mu_L(\{\omega : {}^\circ x_n(\omega) \leq \alpha\}) = lim_{m \to \infty} {}^\circ \nu(\{\omega : x_n(\omega) < \alpha + \tfrac{1}{m}\})$$

$$= lim_{m \to \infty} F(\alpha + \tfrac{1}{m}) = F(\alpha)$$

[28] Given a $*$-finite sequence $(x_m)_{0 \leq k \leq m}$ of random variables on $(X, \mathfrak{A}, \nu)$, we define their distribution function $H : {}^*\mathcal{R}^m \rightarrow^* \mathcal{R}$ by setting $H(\alpha_1, \ldots, \alpha_m) = \nu(\{\omega : x_1(\omega) \leq \alpha_1, \ldots, x_m(\omega) \leq \alpha_m\})$, for an internal tuple $(\alpha_1, \ldots, \alpha_m) \in {}^*\mathcal{R}^m$. Such a function inherits all the properties of standard distribution functions by transfer. When m is finite, H may itself be not only internal, but the transfer of a standard function $H' : \mathcal{R}^m \rightarrow \mathcal{R}$, as is required here. If the sequence $(x_m)_{0 \leq k \leq m}$ is $*$-independent, then H can be written as the $*$-finite product of the distribution functions H_k of each x_k. This follows as H and H_m are $*$-right continuous, and by the definition of $*$-independence.

This means that $G = F|\mathcal{R}$, so $F = {}^*G$, as F was assumed to be the transfer of a standard function. moreover, it implies that each random variable ${}^\circ x_n$ has the same cumulative distribution function G. By Lemma 7.2, $\{{}^\circ x_n\}_{n\in\mathcal{N}}$ is an iid (independent and identically distributed) sequence, ([29]). Using the fact, from transfer, that;

$$E({}^\circ x_n) = \int_{\mathcal{R}} x dG(x) = \int_{{}^*\mathcal{R}} x dF(x)dx = E(x_n)$$

we have $E({}^\circ x_n) = 0$ and, similarly, $E({}^\circ x_n{}^2) = 1$. By the standard Central Limit Theorem, see [15], ([30]), if $\alpha \in \mathcal{R}$ and $\epsilon \in \mathcal{R}_{>0}$ are fixed, then there exists $n_0 \in \mathcal{N}$, such that for all $n > n_0$, with $n \in \mathcal{N}$;

$$|\mu_L(\{\omega : \frac{1}{\sqrt{(n)}} \sum_{k=0}^{n} {}^\circ x_k(\omega) \leq \alpha\}) - \Psi(\alpha)| < \epsilon$$

Now, since the sequence $\{{}^\circ x_k\}_{k\in\mathcal{N}}$ is independent, the cumulative distribution function H of $\sum_{k=0}^{n} {}^\circ x_k$ is G^n, the $n'th$ convolution product of G, ([31]). Thus, for $n > n_0$;

$$|G^n(\sqrt{n}\alpha) - \Psi(\alpha)| < \epsilon$$

As $F =^* G$, we obtain, using the proof of Theorem 2.22(i), that for any $n \in {}^*\mathcal{N} \setminus \mathcal{N}$, $F^n(\sqrt{n}\alpha) \simeq \Psi(\alpha)$. Using the facts that we can express independence in the language $\mathcal{L}_{\mathcal{R}}$, $\{x_k\}_{k\in{}^*\mathcal{N}}$ is $*$-independent and the above footnote, we obtain, by transfer, that F^n is the distribution function of $\sum_{k=0}^{n} x_k$. Hence, for $\alpha \in \mathcal{R}$;

[29] If the random variables have ${}^\circ x_n$ have the same cdf G, then they have the same pdf(probability density function) a.e, this can be shown easily in the discrete case, and the continuous case follows using the Radon-Nikodym theorem applied to the absolutely continuous measure ${}^\circ x_{n*}\mu_L$. The remaining case is obtained using the fact that G has only countably many discontinuities, see the proof of Theorem 3.12

[30] The analytic part of the proof of this result, not covered in the cited text, requires the result proved in Section 5, so, in a sense, we have used entirely non-standard methods to obtain Theorem 7.4.

[31] Technically, this is a slight abuse of the term. If h denotes the pdf corresponding to H, then h is the $n'th$ convolution product of g, where g is the pdf corresponding to G. This follows, at least in the continuous case, from the Inversion Theorem, see Section 5, and the convolution theorem, see [40]. Clearly, we can write H as an analytic expression involving g, and, therefore, H is determined uniquely from a logical expression involving G. We can then transfer this relationship to ${}^*\mathcal{R}$

$\nu(\{\omega : \frac{1}{\sqrt{(n)}} \sum_{k=0}^{n} x_k(\omega) \leq \alpha\}) \simeq \Psi(\alpha)$ $(***)$

Now since Ψ is continuous, both sides of $(***)$ are increasing, and ${}^*\Psi(\alpha) \simeq 1$, ${}^*\Psi(-\alpha) \simeq 0$, for $\alpha > 0$ infinite, we obtain, using a simple adaptation of Theorem 2.24, that $(***)$ holds for all $\alpha \in {}^*\mathcal{R}$.

□

We make the following definition;

Definition 7.5. *Given a probability space $(\Omega, \mathfrak{D}, P)$, we define a Brownian motion or Wiener process to be a function $W : [0,1] \times \Omega \rightarrow \mathcal{R}$, such that;*

(i). W is a stochastic process, that is, for all $t \in [0,1]$, $W_t = W(t, \circ)$ is $\mathfrak{D}$-measurable.

(ii). If $s < t$ belong to $[0,1]$, then $W_t - W_s$ has a normal distribution, with mean 0 and variance $t - s$.

(iii). If $s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n$ belong to $[0,1]$, then;

$\{W_{t_1} - W_{s_1}, \ldots, W_{t_n} - W_{s_n}\}$ is an independent set of random variables.

We define two stochastic processes X_1 and X_2 to be equivalent, if, for all $t \in [0,1]$;

$P(\{\omega : X_1(t,\omega) \neq X_2(t,\omega)\}) = 0$

Remarks 7.6. *A useful consequence of $(ii), (iii)$ is that a Brownian motion is a Gaussian process, in the sense that, for any finite set of times $0 \leq t_1 < \ldots < t_n \leq 1$, the vector $(W_{t_1}, \ldots, W_{t_1})$ has a multivariate Gaussian distribution, see* [39]. *It is a well known fact, due to Norbert Wiener, that there exists a Brownian motion, defined on the probability space $[0,1]$, with Lebesgue measure. In the definition above, we do not require that the paths $W(\circ, \omega)$ are continuous, but it can be shown, that any Brownian motion is equivalent to a Brownian motion in which all the paths are continuous. We will construct a Brownian motion, using non standard methods, in which almost all paths are continuous. It is intuitively clear, but a little tricky to prove rigorously,* *that a.e dP, the Brownian motion path $W(\circ,\omega)$ is not differentiable for any $t\in[0,1]$.*

We now turn to the non-standard construction of Brownian motion.

Definition 7.7. *We let $\eta\in{}^*\mathcal{N}$ and $\Omega=\{-1,1\}^\eta$ be the hyperfinite set consisting of internal sequences of 1's and −1's, indexed by η. We let $\mathfrak{A}$ be the hyperfinite algebra on Ω consisting of all internal subsets of Ω, and ν be the counting measure on $\mathfrak{A}$, defined by $\nu(A)=\frac{Card(A)}{2^\eta}$. We denote by $(\Omega,\mathfrak{D},P)=(\Omega,L(\mathfrak{A},L(\nu)$ the Loeb probability space associated to $(\Omega,\mathfrak{A},\nu)$. Define a $*$-random walk on $(\Omega,\mathfrak{A},\nu)$, by;*

$\chi(t,\omega)=\frac{1}{\sqrt{\eta}}({}^*\sum_{i=1}^{[\eta t]}\omega_i+(\eta t-[\eta t])\omega_{[\eta t]+1})$, [32]

for $t\in{}^*[0,1]$ and $\omega\in\Omega$. Let;

$W(t,\omega)={}^\circ\chi(t,\omega)$

for $(t,\omega)\in[0,1]\times\Omega$.

We now have the following;

Theorem 7.8. *If $\eta\in{}^*\mathcal{N}\setminus\mathcal{N}$, then W is a Brownian motion on $(\Omega,\mathfrak{D},P)$, [33].*

Proof. We show each property in turn;

(i). Fixing $t\in[0,1]$, $\chi(t,\circ)$ is an internal function of ω, hence it is $\mathfrak{A}$-measurable, using Lemma 2.10. Therefore, $W(t,\circ)={}^\circ\chi(t,\circ)$ is $\mathfrak{D}$-measurable by Theorem 3.15(3.8).

(ii). Fixing $s<t$ in $[0,1]$, we have that;

$P(\{\omega:W(t,\omega)-W(s,\omega)\leq\alpha\})$

---

[32] Here, [ ] denotes the integer part, that is, for $r\in{}^*\mathcal{R}$, $[r]$ is the largest hyperfinite integer n for which $n\leq r$. χ is obtained, by transfer, from finite random walks, with step size $\frac{1}{\sqrt{n}}$ on the spaces Ω_n of coin tossings of length n. Observe that the remainder term ensures the paths $\chi(\omega,\circ)$ are $*$-continuous.

[33] Technically, W can take infinite values, but $\{\omega:\exists tW(t,\omega)\in\{+\infty,-\infty\}\}$ has Loeb measure zero, this will follow from an even stronger result in the next Theorem.

$= P(\{\omega : {}^\circ\chi(t,\omega) - {}^\circ\chi(s,\omega) \leq \alpha\})$

$= P(\{\omega : \frac{1}{\sqrt{\eta}}({}^\circ \sum_{i=[\eta s]}^{[\eta t]}\omega_i) \leq \alpha\})$ $(*)$

$= lim_{n\rightarrow\infty}{}^\circ\nu(\{\omega : \frac{1}{\sqrt{\lambda}}\sum_{i=[\eta s]}^{[\eta t]}\omega_i \leq \sqrt{\frac{\eta}{\lambda}}(\alpha+\frac{1}{n})\})$, $(**)$, $\lambda = [\eta t] - [\eta s]$

$= lim_{n\rightarrow\infty}{}^\circ({}^*\Psi)(\sqrt{\frac{\eta}{\lambda}}(\alpha+\frac{1}{n}))$ $(***)$

$= lim_{n\rightarrow\infty}\Psi({}^\circ(\sqrt{\frac{\eta}{\lambda}}(\alpha+\frac{1}{n})))$

$= lim_{n\rightarrow\infty}\Psi(\frac{\alpha+\frac{1}{n}}{\sqrt{t-s}})$ $(****)$

$= \Psi(\frac{\alpha}{\sqrt{t-s}})$ $(\dagger)$

We justify the steps in the calculation. $(*)$ follows from the fact that the $*$-finite random walk has infinitesimal step size, so, when taking the standard part, we can ignore the remainder terms. $(**)$ uses a similar argument to Lemma 7.2, and a straightforward rearrangement of terms. $(***)$ uses Theorem 7.4. In order to apply the theorem, we need to check that $\{\omega_i\}_{[\eta s]\leq i\leq[\eta t]}$ is a $*$-independent sequence of random variables, with common standard distribution function F, and with $E(\omega_i) = 0$, $E(\omega_i^2) = 1$. However, these facts all follow, by transfer, from the corresponding results for finite random walks with step size 1 on Ω_n, for finite n, see above footnote. The distribution function F on ${}^*\mathcal{R}$ can be written down explicitly as;

$F(x) = 0$, for $x < -1$, $F(x) = \frac{1}{2}$, for $-1 \leq x < 1$, $F(x) = 1$, for $x \geq 1$

so $E(\omega_i) = \frac{1}{2}.1 + \frac{1}{2}.-1 = 0$ and $E(\omega_i^2) = \frac{1}{2}.1 + \frac{1}{2}.1 = 1$ as required. $(****)$ follows from the simple equality ${}^\circ\sqrt{\frac{\eta}{\lambda}} = \frac{1}{\sqrt{t-s}}$, ([34]). The rest of the calculation just uses continuity of Ψ, and Theorem 2.25.

By $(\dagger)$, $P(\{\omega : W(t,\omega) - W(s,\omega) \leq \alpha\sqrt{t-s}\}) = \Psi(\alpha)$, so $W_t - W_s$ has a normal distribution with mean zero and variance $t-s$, as required.

(iii). Suppose that $s_1 < t_1 \leq s_2 < t_2 \leq \ldots \leq s_n < t_n$ belong to $[0,1]$. Then;

$\{\chi(t_1,\circ) - \chi(s_1+\frac{1}{\eta},\circ), \ldots, \chi(t_n,\circ) - \chi(s_n+\frac{1}{\eta},\circ)\}$

[34]We have that ${}^\circ(\frac{\eta}{\lambda}) = {}^\circ(\frac{\eta}{[\eta t]-[\eta s]}) = {}^\circ(\frac{\eta}{\eta t - c - \eta s + d}) = {}^\circ(\frac{1}{t-s+\frac{d-c}{\eta}}) = \frac{1}{t-s}$, as $(0 \leq c, d < 1)$

is $*$-independent, hence S-independent, ([35]). By Lemma 7.2;

$\{W_{t_1} - W_{s_1}, \ldots, W_{t_n} - W_{s_n}\}$

is independent, as clearly $^\circ\chi(s_i + \frac{1}{\eta}, \omega) = {}^\circ\chi(s_i, \omega)$ for any ω and i.

□

Theorem 7.9. *There exists a set $\Omega' \subset \Omega$, with $P(\Omega') = 1$, such that $W(\circ, \omega)$ is continuous and finite for almost all $\omega \in \Omega'$, in fact $\chi(\circ, \omega)$ is near standard in $^*C[0,1]$ for all $\omega \in \Omega'$.*

Proof. This is an important technical result, but I did not make any clarifications which are not already in [1]. □

Definition 7.10. *Wiener measure is defined as the unique Borel measure μ on $C[0,1]$ such that;*

(i). $\mu(\{f : f(t) < \alpha\}) = \Psi(\frac{\alpha}{\sqrt{t}})$

(ii). If $s_1 < t_1 \leq \ldots \leq s_n < t_n$ belong to $[0,1]$, then the random variables;

$\{f(t_1) - f(s_1), \ldots, f(t_n) - f(s_n)\}$

are independent.

Lemma 7.11. *Given $(\Omega, \mathfrak{D}, P)$ as above, let $\mathfrak{C}$ be the σ-algebra on $C[0,1]$ defined by;*

$E \in \mathfrak{C}$ *iff* $\{\omega : W(\circ, \omega) \in E\} \in \mathfrak{D}$

and let P' be the probability measure defined by $P'(E) = P(\{\omega : W(\circ, \omega) \in E\}$, then;

$\mathfrak{C}$ contains all the Borel sets on $C[0,1]$ and, if $\eta \in {}^*\mathcal{N} \setminus \mathcal{N}$, then $(C[0,1], \mathfrak{C}, P')$ is an extension of Wiener measure.

Proof. This is a straightforward consequence of the properties of the Brownian motion W, constructed in Theorem 7.8. □

[35]The extra $\frac{1}{\eta}$ just ensures that the $*$-finite sums do not overlap, in the worst case scenario, when $s_{i+1} = t_i$, for some $1 \leq i \leq n-1$. This is a straightforward calculation left to the reader.

Theorem 7.12. *Let P'_η be the probability measure given by the previous theorem, depending on η, then;*

$\{P'_\eta\}_{\eta\in\mathcal{N}}$ *converges weakly to P', ([36]).*

Proof. The proof is mainly an exercise in Loeb integration theory, using Definition 7.7 and Theorem 7.8. □

We now develop the theory of stochastic calculus. We first recall some of the standard theory.

Definition 7.13. *Let W be a Brownian motion on a complete probability space $(\Omega, \mathfrak{D}, P)$, then, we define a filtration $\{\mathfrak{D}_t\}_{t\in[0,1]}$ to be a collection of sub σ-algebras of $\mathfrak{D}$, such that;*

(i). $\mathfrak{D}_t \supset \mathfrak{D}_s$ *if* $s < t$.

(ii). $W(t, \circ)$ *is* $\mathfrak{D}_t$ *measurable, for all* $t \in [0, 1]$.

(iii). If $t \leq s_1 < t_1 \leq \ldots \leq s_n < t_n$ *belong to* $[0, 1]$*, then* $\mathfrak{D}_t$ *is independent of the σ-algebra generated by;*

$\{W(t_1, \circ) - W(s_1, \circ), \ldots, W(t_n, \circ) - W(s_n, \circ)\}$

Let $([0, 1], \mathfrak{B}, \mu)$ *be the* complete *probability space on the interval* $[0, 1]$*, where* $\mathfrak{B}$ *is the completion of the Borel field with respect to Lebesgue measure* μ*, and let* $([0, 1] \times \Omega, \mathfrak{B} \times \mathfrak{D}, \mu \times P)$ *be the* complete *product of* $([0, 1], \mathfrak{B}, \mu)$ *and* $(\Omega, \mathfrak{D}, P)$*. Then, we define* $\mathcal{G}_0$ *to be the set of all functions* g *such that;*

(i). $g \in L^2([0, 1] \times \Omega, \mathfrak{B} \times \mathfrak{D}, \mu \times P)$.

(ii). $g|_{[0,t]\times\Omega}$ *is* $\mathfrak{B} \times \mathfrak{D}_t$*-measurable,([37]), for all* $t \in [0, 1]$.

We define $\mathcal{G}'_0 \subset \mathcal{G}_0$ *to be the set of all functions* $g \in \mathcal{G}_0$*, satisfying the additional assumption;*

[36]See Section 6, for a definition of weak convergence. This is a special case of Donsker's Theorem, see [30]. In fact, one can deduce Donsker's theorem from this result, as is done in [22].

[37]Here, we are just taking the product, and not passing to the completion

(iii) There exist $0 = t_0 < t_1 < \ldots < t_n = 1$ *belonging to* $[0, 1]$*, such that;*

For all $\omega \in \Omega$*,* $g(t, \omega) = g(t_i, \omega)$*, for* $t \in [t_i, t_{i+1})$*,* $0 \leq i \leq n - 1$

If $g \in \mathcal{G}'_0$*, and* $t \in [0, 1]$*, we define the standard stochastic integral by;*

$\int_0^t g(\tau, \omega) dW(\tau, \omega)$

$= \sum_{i=0}^{k-1} g(t_i, \omega)(W(t_{i+1}, \omega) - W(t_i, \omega)) + g(t_k, \omega)(W(t, \omega) - W(t_k, \omega))$

where $k = max\{i : 0 \leq i \leq n - 1, t_i \leq t\}$.

For $g' \in \mathcal{G}'_0$*, and* $t \in [0, 1]$*, we have that* $|| \int_0^t g' dW||_2 = ||g'||_2$*, see* [39]*. Using the fact that* $\mathcal{G}'_0$ *is dense in* $\mathcal{G}_0$*, we define the stochastic integral* $\int_0^t g dW$*, for* $g \in \mathcal{G}_0$*, by extension, ([38]). Observe that the integral is only defined up to stochastic equivalence.*

We now turn to a non standard version of the stochastic integral. In order to define this, it is convenient to modify the Loeb probability space $(\Omega, \mathfrak{D}, P) = (\Omega, L(\mathfrak{A}), L(\nu))$, slightly, and its associated Brownian motion;

Definition 7.14. *We now let* $\Omega = \{-1, 1\}^{2\eta+1}$ *consist of internal tuples of* 1*'s and* -1*'s, indexed by* $I = \{i \in {}^*\mathcal{N} : -\eta \leq i \leq \eta\}$*. As before,* $\mathfrak{A}$ *is the hyperfinite* ${}^*\sigma$*-algebra, of internal subsets of* Ω*,* ν *is counting measure,* $(\Omega, \mathfrak{D}, P) = (\Omega, L(\mathfrak{A}), L(\nu))$ *is the associated Loeb space, and* χ *and* W *are as in Definition 7.7, ([39]). We now let* $\Omega'' = \{\omega \in \Omega : \omega_i \in \Omega', \forall i \geq 1\}$*, where* Ω' *is given by Theorem 7.9, so* $P(\Omega'') = 1$*. For* $0 \leq i \leq \eta$*, we let* $\sim_i$ *be the equivalence relation on* Ω*, defined by;*

$\omega \sim_i \omega'$ *iff* $\omega_j = \omega'_j$ *for all* $j \leq i$

For $0 \leq i \leq \eta$*, we let* $\mathfrak{A}_i$ *be the algebra generated by the partition of* Ω *into equivalence classes with respect to* $\sim_i$*, and* $\mathfrak{H}_i$ *be the external* σ*-algebra consisting of unions of these equivalence classes. We let* $L(\mathfrak{A}_i)$ *denote* $\mathfrak{H}_i \cap \sigma(\mathfrak{A}_i)^{comp}$*, where* $\sigma(\mathfrak{A}_i)^{comp}$ *is the completion of the*

[38] This extension is known as Ito's isometry

[39] It is easy to see that W is still a Brownian motion on $(\Omega, \mathfrak{D}, P)$

σ-algebra generated by $\mathfrak{A}_i$. For $t \in [0,1]$, we let $\mathfrak{D}_t$ be the σ-algebra generated by $\bigcap_{i\geq 0,\frac{i}{\eta}\simeq t} L(\mathfrak{A}_i)$ and $W(t,\circ)$.

Lemma 7.15. *$\{\mathfrak{D}_t\}_{t\in[0,1]}$, as in Definition 7.14, is a filtration in the sense of Definition 7.13.*

Proof. We check each of the conditions;

(i). Observe that, if $i<j$, $\sim_j$ refines $\sim_i$, so $\sigma(\mathfrak{A}_i)^{comp} \subset \sigma(\mathfrak{A}_j)^{comp}$, $\mathfrak{H}_i \subset \mathfrak{H}_j$ and $L(\mathfrak{A}_i) \subset L(\mathfrak{A}_j)$. Clearly, if $s<t$, and $\frac{i}{\eta}\simeq s$, then if $\frac{j}{\eta}\simeq t$, we must have that $i<j$, so $\bigcap_{i\geq 0,\frac{i}{\eta}\simeq s} L(\mathfrak{A}_i) \subset \bigcap_{i\geq 0,\frac{i}{\eta}\simeq t} L(\mathfrak{A}_i) \subset \mathfrak{D}_t$. Now observe that ω_i is measurable with respect to $\mathfrak{A}_j$, for any $j\geq i$, therefore, $\chi(s,\circ)$ is measurable with respect to $\mathfrak{A}_j$, for any $j\geq[\eta s]+1$. By Lemma 3.15(3.8), $W(s,\circ)$ is measurable with respect to $L(\mathfrak{A}_j)$, for any $j\geq[\eta s]+1$. If $s<t$, and $\frac{j}{\eta}\simeq t$, then $j\geq[\eta s]+1$, ([40]), so $W(s,\circ)$ is measurable with respect to $\mathfrak{D}_t$. Therefore, $\mathfrak{D}_s \subset \mathfrak{D}_t$.

(ii). Clear by definition.

(iii). By the definition of W, we can choose $i\geq 0$ with $\frac{i}{\eta}\simeq t$ such that $i<[\eta s_1]+1$ and $B\in L(\mathfrak{A}_i)$. Observe that by the definition of $\sim_i$ and the observation in (i), $\mathfrak{A}_i$ is exactly the ${}^*\sigma$-algebra generated by the random variables ω_j, for $j\leq i$. As in Theorem 7.8(iii), it follows, by transfer, that $\mathfrak{A}_i$ is independent, with respect to ν, of the ${}^*\sigma$-algebra generated by $\chi(t_1,\circ)-\chi(s_1,\circ)$. It follows easily, using Theorem 3.15(3.4) to approximate, and Theorem 3.15(3.8), that $L(\mathfrak{A}_i)$ is independent of the σ algebra generated by $W(t_1,\circ)-W(s_1,\circ)$, with respect to P. By a similar argument, using transfer, $L(\mathfrak{A}_i)$ is independent of the σ-algebra $\mathfrak{Q}$ generated by $\{W(t_1,\circ)-W(s_1,\circ),\ldots,W(t_n,\circ)-W(s_n,\circ)\}$, so $\bigcap_{i\geq 0,\frac{i}{\eta}\simeq t} L(\mathfrak{A}_i)$ is independent of this σ-algebra. Now, using Theorem 7.9, and the fact that almost sure convergence implies convergence in probability, see [39], if $t>0$, and $B\in\sigma(W(t,\circ))$, we can find sets $B_{(0\leq t'\leq t)}$ with $B_{t'}\in\sigma(W(t',\circ))$ such that for all $D\in\mathfrak{D}$, $P(D\cap B_t)=lim_{t'\to t}P(D\cap B_{t'})$, $(*)$. Then, repeating the argument above with $D\in\bigcap_{i\geq 0,\frac{i}{\eta}\simeq t} L(\mathfrak{A}_i)$, and using (i), we obtain that $D\cap B_{t'}$ is independent of $\mathfrak{Q}$, for $t'<t$. Using $(*)$ and the definition of independence, we obtain $D\cap B_t$ is independent of $\mathfrak{Q}$. Now using a π-systems argument, see [44], we obtain the result.

□

[40] You need η to be infinite, here.

We now require the following;

Definition 7.16. *Let* $(S, \mathfrak{B}, \mu)$ *be the* complete *probability space on the interval* $[0, 1]$, *with* 0 *and* 1 *identified, where* $\mathfrak{B}$ *is the completion of the Borel field with respect to Lebesgue measure* μ, *and let* $(S\times\Omega, \mathfrak{B}\times\mathfrak{D}, \mu\times P)$ *be the* complete *product of* $(S, \mathfrak{B}, \mu)$ *and* $(\Omega, \mathfrak{D}, P)$.

Let $\eta \in {}^*\mathcal{N} \setminus \mathcal{N}$ *be as above, and* $Y = {}^*[0, 1]$, *with* 0 *and* 1 *identified. We let* $\mathfrak{C}$ *denote the* ${}^*\sigma$*-algebra generated by all internal unions of intervals of the form* $[\frac{i}{\eta}, \frac{i+1}{\eta})$, *for* $0 \leq i < \eta$, *and* $i \in {}^*\mathcal{N}$. λ *is the* **-additive probability measure defined by* $\lambda([\frac{i}{\eta}, \frac{i+1}{\eta})) = \frac{1}{\eta}$, *and* $(Y, L(\mathfrak{C}), L(\lambda))$ *is the associated Loeb space. We let* $st : Y \rightarrow S$ *be the standard part mapping.*

We recall the following special case of Theorem 5.7, see also Lemma 6.4;

Theorem 7.17. $st : (Y, L(\mathfrak{C}), L(\lambda)) \rightarrow (S, \mathfrak{B}, \mu)$ *is measurable and measure preserving. If* $g : S \rightarrow \mathcal{R}$ *is integrable with respect to* μ, *then* $g \circ st$ *is integrable with respect to* $L(\lambda)$ *and;*

$\int_{st^{-1}(B)}(g \circ st)dL(\lambda) = \int_B g d\mu$

for all $B \in \mathfrak{B}$.

Definition 7.18. *We define* $\mathcal{G}$ *to be the set of all functions* g *such that;*

(i). $g \in L^2([0, 1] \times \Omega, \mathfrak{B} \times \mathfrak{D}, \mu \times P)$.

(ii). For all $t \in [0, 1]$, $g(t, \circ)$ *is* $\mathfrak{D}_t$*-measurable,* ([41]).

Suppose $g : [0, 1] \times \Omega \rightarrow \mathcal{R} \cup \{+\infty, -\infty\}$, *then we say* f *is a p-lifting of* g *if;*

$(i)'$. $f \in SL^p(Y \times \Omega, \mathfrak{C} \times \mathfrak{A}, \lambda \times \nu)$

$(ii)'$. ${}^\circ f(t, \omega) = g({}^\circ t, \omega)$ *a.e* $L(\lambda \times \nu)$.

$(iii)'$. *For all* $t \in {}^*[0, 1)$, $f(t, \circ)$ *is* $\mathfrak{A}_{[\eta t]}$*-measurable.*

[41]Observe that $\mathcal{G}_0 \subset \mathcal{G}$, this follows from Theorem 8.2 of [37].

We let $\mathcal{F}$ be the class of g having 2-liftings.

Theorem 7.19. *Suppose that $g \in L^p([0,1] \times \Omega)$, and, for all $t \in [0,1]$, $g(t, \circ)$ is $\mathfrak{D}_t$-measurable, then g has a p-lifting f. In particular, $\mathcal{G} \subset \mathcal{F}$.*

Proof. Let $\mathfrak{D}'_t = \bigcap_{i \geq 0, \frac{i}{\eta} \simeq t} L(\mathfrak{A}_i)$. We first modify g on a set of P-measure zero, to find g' such that $g'(t, \circ)$ is $\mathfrak{D}'_t$-measurable. Define an external equivalence relation $\sim_t$ on Ω by $\omega \sim_t \omega'$ iff $\omega_j = \omega'_j$ for all $j \leq 0$ and $^{\circ}(\frac{j}{\eta}) < t$. Define g' by;

$g'(t, \omega) = g(t, \omega')$ if there exists $\omega' \in \Omega'$ with $\omega \sim_t \omega'$

$g'(t, \omega) = 0$ otherwise.

In order to see that g' is well defined, suppose that ω', ω'' belong to Ω' and $\omega' \sim_t \omega''$. If $s < t$, then as $\chi(s, \circ)$ is defined in terms of ω_j, for $j \leq [\eta s] + 1$, we clearly have that $\chi(s, \omega') = \chi(s, \omega'')$, so $W(s, \omega') = W(s, \omega'')$, $(*)$. As $W(\circ, \omega)$ is continuous for $\omega \in \Omega'$, by Theorem 7.9, it follows that $W(t, \omega') = W(t, \omega'')$. Now, considering the internal set $\{j : \omega'_j = \omega''_j\}$, it follows, using overflow, that there exists j with $j\eta \simeq t$ such that $\omega' \sim_j \omega''$, therefore, as $L(\mathfrak{A}_j) \subset \mathfrak{H}_j$, we have that for every set $B \in \mathfrak{D}'_t$, $\omega' \in B$ iff $\omega'' \in B$. By considering the probability measures $\delta_{\omega'}$ and $\delta_{\omega''}$, and using $(*)$, it follows, using Lemma 1.6(b) in [44], that for every $B \in \mathfrak{D}_t$, $\omega \in B'$ iff $\omega'' \in B$. As $g(t, \circ)$ is $\mathfrak{D}_t$-measurable, by considering $g_t^{-1}(g_t(\omega'))$, it follows that $g(t, \omega') = g(t, \omega'')$, as required. If $(\frac{i}{\eta}) \sim t$, and $\omega' \sim_i \omega''$, then either there exists $\omega''' \in \Omega'$ with $\omega''' \sim_t \omega'$, in which case we have that $g'(t, \omega') = g'(t, \omega'') = g'(t, \omega''')$ by definition, or $g'(t, \omega') = g'(t, \omega'') = 0$, so $g'(t, \circ)$ is constant on the equivalence classes generating $\mathfrak{A}_i$, for $(\frac{i}{\eta}) \sim t$, $(**)$. Now clearly $g'(t, \circ)$ is $L(\mathfrak{A})$-measurable, (check the definition of g' can be written in terms of a countable intersection of sets belonging to $\mathfrak{A}$). Therefore, by Theorem 7.3(i) and $(**)$, $g'(t, \circ)$ is $L(\mathfrak{A}_i)$-measurable, for $(\frac{i}{\eta}) \sim t$, so $g'(t, \circ)$ is $\mathfrak{D}'_t$-measurable. Now, by definition, for any $t \in [0,1]$, g_t and g'_t agree on Ω', so, checking the requirement (ii) of a lift, and using the fact that $L(\lambda \times \nu)(Y \times \Omega') = 1$, we can replace g by g', and so assume g is $\mathfrak{D}'_t$-measurable.

Now let $g_1(t, \omega) = g(^{\circ}t, \omega)$, for $(t, \omega) \in Y \times \Omega$, so g_1 is a lift of g to $Y \times \Omega$. By a simple extension of Theorem 7.17, using the fact that $g \in L^p([0,1] \times \Omega, \mathfrak{B} \times L(\mathfrak{A}), \mu \times L(\nu))$, we have $g_1 \in L^p(Y \times \Omega, L(\mathfrak{C}) \times L(\mathfrak{A}, L(\lambda) \times L(\nu))$. By Lemma 3.38, see also the proof of

Lemma 3.15(3.12), $g_1 \in L^p(Y\times\Omega, L(\mathfrak{C}\times\mathfrak{A}), L(\lambda\times\nu))$. Now applying Theorem 3.34(ii), we can find $f_1 \in SL^p(Y\times\Omega, \mathfrak{C}\times\mathfrak{A}, \lambda\times\nu)$, with;

$$^\circ f_1(t,\omega) = g_1(t,\omega) = g(^\circ t,\omega) \text{ a.e } dL(\lambda\times\nu)\ (***)$$

Clearly, f_1 satisfies the requirements $(i)', (ii)'$. In order to obtain $(iii)'$, we need to modify f_1 again slightly. We let $\mathfrak{C}'$ be the $^*\sigma$-subalgebra of $\mathfrak{C}\times\mathfrak{A}$ generated by the hyperfinite partition $P = \{[\frac{i}{\eta}, \frac{i+1}{\eta})\times A : 0 \leq i < \eta, A \in \mathfrak{A}_i\}$. Let $f = E[f_1|\mathfrak{C}']$, see Definition 7.1. By Theorem 7.3(ii), $f \in SL^p(Y\times\Omega, \mathfrak{C}', \lambda\times\nu)$, (so $(i)'$ holds), and $^\circ f = E(^\circ f_1|L(\mathfrak{C}')$. As $^\circ f_1 = g_1$ a.e $dL(\lambda\times\nu)$, by $(***)$, and $L(\mathfrak{C}') \subset L(\mathfrak{C}\times\mathfrak{A})$, it follows, using [44] Theorem 9.29(c), that $E(^\circ f_1|L(\mathfrak{C}')) = E(g_1|L(\mathfrak{C}'))$, so $^\circ f = E(g_1|L(\mathfrak{C}'))$, $(****)$. We claim that g_1 is constant on elements of P, $(\dagger)$. For if $t \in [\frac{i}{\eta}, \frac{i+1}{\eta})$ and $A \in \mathfrak{A}_i$, then $g_{1,t}|_A = g_s|_A$, using $(***)$, where $s = {^\circ}\frac{i}{\eta}$. By the definition of $\mathfrak{D}'_s$, and the fact that g_s is $\mathfrak{D}'_s$-measurable, g_s is constant on A, so $(\dagger)$ holds. Now applying Theorem 7.3(i), g_1 is $L(\mathfrak{C}')$-measurable. Hence, $E(g_1|L(\mathfrak{C}')) = g_1$ a.e $dL(\lambda\times\nu)$, so, by $(****)$, $^\circ f = g_1$ a.e d $L(\lambda\times\nu)$, (so $(ii)'$ holds). Now, if $t \in Y$, we can choose i, such that $t \in [\frac{i}{\eta}, \frac{i+1}{\eta})$, so $i = [\eta t]$. As f is constant on elements of P, and P is a partition, $f_t = f_{\frac{i}{\eta}}$, and, clearly $f_{\frac{i}{\eta}}$ is measurable with respect to $\mathfrak{A}_i$, by definition of P. Therefore, $f(t,\circ)$ is $\mathfrak{A}_{[\eta t]}$-measurable, (so $(iii)'$ holds.) ([42]). $\square$

Definition 7.20. *Suppose $g \in \mathcal{F}$, and f is a 2-lifting of g. Then, for $t \in [0,1]$, we define;*

$$I(t,\omega) = \int_0^t g(\tau,\omega)dW(\tau,\omega) = {^\circ}\int_0^t f(\tau,\omega)d\chi(\tau,\omega)$$

where the final integral is a *-Stieltjes integral.

Remarks 7.21. *As $f(\circ,\omega)$ is $\mathfrak{C}$-measurable, it is constant on each interval $[\frac{i}{\eta}, \frac{i+1}{\eta})$, for $0 \leq i < \eta$. $\chi(\circ,\omega)$ defines a *-continuous internal function on $^*[0,1]$. Hence, we can form the hyperfinite sum, see Definition 2.19;*

$$F(t,\omega) = \int_0^t f(\tau,\omega)d\chi(\tau,\omega)$$

$$= \sum_{i=0}^{[\eta t]-1} f(\tfrac{i}{\eta},\omega)[\chi(\tfrac{i+1}{\eta},\omega) - \chi(\tfrac{i}{\eta},\omega)] + f(\tfrac{[\eta t]}{\eta},\omega)[\chi(t,\omega) - \chi(\tfrac{[\eta t]}{\eta},\omega)]$$

[42] f is referred to in [12] as a 2-legged lifting of g, not to be confused with the 2 in L^2.

This is defined for all $t \in {}^*[0,1]$, if we adopt the convention that $f(1,\omega)=0$. Clearly, this is the transfer of the standard Stieltjes integral, defined on a finite partition of $[0,1]$. Observe the striking similarity between the definition of the *-Stieltjes integral in terms of the increments of a random walk and the standard definition of the stochastic integral in terms of the increments of Brownian motion.

Theorem 7.22. *Definition 7.20 is independent of the choice of lift f for g. Moreover, it coincides with Ito's definition of the stochastic integral on $\mathcal{G}_0$.*

Proof. It is sufficient to consider the case $t=1$. Let $F(\omega)=\int_0^1 f(t,\omega)d\chi(t,\omega)$. It is easy to check, using the definition of the *-Stieltjes integral, and Lemma 2.10, that F is $\mathfrak{A}$-measurable. Hence,the internal integral of F^2 with respect to ν belongs to ${}^*\mathcal{R}$. We then compute;

$$^\circ\int_\Omega F(\omega)^2 d\nu = {}^\circ\int_\Omega (\sum_{i=0}^{\eta-1} f(\tfrac{i}{\eta},\omega)\tfrac{\omega_{i+1}}{\sqrt{\eta}})^2 d\nu \ (*)$$

$$= {}^\circ\int_\Omega \sum_{i=0}^{\eta-1} \frac{f^2(\frac{i}{\eta},\omega)}{\eta} d\nu + 2^\circ\int_\Omega \sum_{0\leq i<j\leq \eta-1} \frac{f(\frac{i}{\eta},\omega)f(\frac{j}{\eta},\omega)\omega_{i+1}\omega_{j+1}}{\eta} d\nu$$

$$= {}^\circ\int_\Omega\int_Y f^2(t,\omega)d\lambda d\nu + 2^\circ\sum_{0\leq i<j\leq\eta-1}(\int_\Omega \omega_{j+1}d\nu)(\int_\Omega \frac{f(\frac{i}{\eta},\omega)f(\frac{j}{\eta},\omega)}{\eta}\omega_{i+1}d\nu)$$

$(**)$

Here, $(*)$ follows from the definition of the *-Stieltjes integral, and the fact that the increments of the *-random walk are given by the random variables $\{\frac{\omega_i}{\sqrt{\eta}}\}_{1\leq i\leq\eta}$. In $(**)$, we have used Remarks 3.10, to obtain the integral over Y, and *-independence of the random variables ω_j with the algebras $\mathfrak{A}_i$, for $i<j$. Observe that this is why Definition 7.18(iii)' of a lifting is required. Now, as $E(\omega_{j+1})=\int_\Omega \omega_{j+1}d\nu=0$, for $0\leq j\leq\eta-1$, we obtain from $(**)$, Theorem 3.16, the transfer of Fubini's theorem, see [37] or [42], and Definition 7.18(i)', that;

$$\int_\Omega {}^\circ F^2(\omega)dL(\nu) \leq {}^\circ\int_\Omega F^2(\omega)d\nu = ||f||_2^2$$

Now, if f_1 and f_2 are different 2-liftings of g, with corresponding *-Stieltjes integrals F_1 and F_2, we can repeat the above argument to obtain;

$$\int_\Omega |{}^\circ F_1 - {}^\circ F_2|^2 dL(\nu) \leq ||f_1-f_2||_2^2 = 0 \ (***)$$

where, we have, finally, used Definition 7.18(i)'(ii)' and Theorem 3.34(i). It follows that ${}^\circ F_1 = {}^\circ F_2$ a.e $dL(\nu)$. Therefore, $I(t,\omega)$, given in Definition 7.20, is well defined up to stochastic equivalence, as required.

Now suppose that $g \in \mathcal{G}'_0$. Then there exist $0 = t_0 < t_1 < \ldots < t_n = 1$, with $g(t,\omega) = g(t_i,\omega)$, for all $t \in [t_i, t_{i+1})$, with $0 \leq i \leq n-1$. Choose $\{t'_0, \ldots, t'_{n-1}\}$ in Y with ${}^\circ t'_i = t_i$ and $\eta t'_i \in {}^*\mathcal{N}$, for $0 \leq i \leq n-1$. Then we claim that there exists a 2-lift f of g with $f(t,\omega) = f(t'_i,\omega)$, for all $t \in *[t'_i, t_{i+1})$, (†). In order to see this, as in Theorem 7.19, modify g to g', so g'_t is $\mathfrak{D}'_t$-measurable, and $g = g'$ on Ω'. We have that $g'_{t_i} \in L^2(\Omega, L(\mathfrak{A}_{[\eta t'_i]}), L(\nu))$, hence, by Theorem 3.34, we can find $f_{t'_i} \in SL^2(\Omega, \mathfrak{A}_{[\eta t'_i]}, \nu)$, (††), with ${}^\circ f_{t'_i} = g'_{t_i} = g_{t_i}$ a.e dP. Let $V \subset \Omega' \subset \Omega$ be a set in $\mathfrak{D}$ for which this holds. Now define f on $Y \times \Omega$, by setting $f(t,\omega) = f_{t'_i}(\omega)$, for $t \in {}^*[t'_i, t'_{i+1})$, where $t'_n = 1$. Then, as the intervals ${}^*[t'_i, t'_{i+1})$ belong to $\mathfrak{C}$, we have that f is $\mathfrak{C} \times \mathfrak{A}$-measurable. We check that f satisfies the conditions $(i)'$,$(ii)'$ and $(iii)'$ of a 2-lift of g, in which case, by definition of f, (†) is shown. For condition $(ii)'$, if $U_i = [t'_i, t'_{i+1}) \cap st^{-1}[t_i, t_{i+1})$, so $P(U_i) = t_{i+1} - t_i$, and $(t,\omega) \in U_i \times V$, then;

$${}^\circ f(t,\omega) = {}^\circ f_{t'_i}(\omega) = g_{t_i}(\omega) = g({}^\circ t, \omega)$$

For condition $(i)'$, we use Lemma 3.19, if there exists an infinite K for which $\int_{|f|^2 > K} |f|^2 d(\lambda \times \nu) > \epsilon$, where ϵ is standard, then, without loss of generality, we can find t'_i, such that $\int_{(|f|^2 > K) \cap [t'_i, t'_{i+1})} |f|^2 d(\lambda \times \nu) > \epsilon$. By definition of f and, using the transfer of Fubini's theorem, we have that $\int_{|f_{t'_i}|^2 > K} |f_{t'_i}|^2 d\nu > \frac{\epsilon}{t_{i+1} - t_i}$, which contradicts (††). For condition $(iii)'$, if $t \in {}^*[t_i, t_{i+1})$, then $f_t = f_{t'_i}$ is $\mathfrak{A}_{[\eta t'_i]}$-measurable. As $[\eta t] \geq [\eta t'_i]$, $\mathfrak{A}_{[\eta t]}$ refines $\mathfrak{A}_{[\eta t'_i]}$, so f_t is $\mathfrak{A}_{[\eta t]}$-measurable. Hence, (†) is shown.

Now given such a lifting f of $g \in \mathcal{G}'_0$, we have, for $\omega \in V$;

$$F(\omega) = \int_0^1 f(t,\omega) d\chi(t,\omega) = \sum_{i=0}^{n-1} f(t'_i,\omega)(\chi(t'_{i+1},\omega) - \chi(t'_i,\omega)) \ (\dagger\dagger\dagger)$$

$$\simeq \sum_{i=0}^{n-1} g(t_i,\omega)(W(t_{i+1},\omega) - W(t_i,\omega)) \ (\sharp)$$

For (†††), we have used the fact that f_ω is constant on ${}^*[t_i, t_{i+1})$, so all the terms in the *-Stieltjes integral, involving χ, cancel, except at the endpoints of the interval. For (♯) we have used the definition of W, Theorem 7.9, and the property $(ii)'$ of f. Hence, ${}^\circ F$ is the standard stochastic integral of g. Now, as we observed above, if $g \in \mathcal{G}_0$, then $||{}^\circ F||_2 \leq ||g||_2$. If I_t denotes the non-standard stochastic integral, and I'_t denotes the standard stochastic integral, then both define continuous linear operators on $\mathcal{G}_0$, which agree on the dense subset $\mathcal{G}'_0$. Hence, they must coincide on $\mathcal{G}_0$. In particular, Ito's isometry holds for the non-standard stochastic integral, I_t, restricted to the class $\mathcal{G}_0$.

□

We illustrate this method with the following example, which does not require Ito's Theorem;

Example 7.23. Let $g(t,\omega) = W(t,\omega)$, for $t \in [0,1]$, and let;

$f(t,\omega) = \frac{1}{\sqrt{\eta}} \sum_{i=1}^{[\eta t]} \omega_i = \chi(\frac{[\eta t]}{\eta}, \omega)$, for $t \in Y$

We claim that f is a 2-lifting of g, so we check all the conditions in Definition 7.18;

$(ii)'$. If $t \in Y$, then, $\frac{[\eta t]}{\eta} \simeq {}^\circ t$, hence, by Theorem 7.9, for $\omega \in \Omega'$;

$f(t,\omega) = \chi(\frac{[\eta t]}{t}, \omega) \simeq \chi({}^\circ t, \omega)$

and

${}^\circ f(t,\omega) = {}^\circ \chi({}^\circ t, \omega) = W({}^\circ t, \omega)$

$(iii)'$. If $\frac{i}{\eta} \leq t < \frac{i+1}{\eta}$, with $i \in {}^*\mathcal{N}$, then $[\eta t] = i$, so f is constant on the elements of the partition P generating $\mathfrak{C}$. As each f_t is $\mathfrak{A}$-measurable, f must be $\mathfrak{C} \times \mathfrak{A}$-measurable. Moreover, for all $t \in {}^*[0,1)$, as ω_i is $\mathfrak{A}_i$-measurable and $\mathfrak{A}_i \subset \mathfrak{A}_j$ if $i < j$, we have that f_t is $\mathfrak{A}_{[\eta t]}$-measurable.

$(i)'$. Let $h(t,\omega) = W({}^\circ t, \omega)$, then, as $W \in L^2([0,1] \times \Omega, \mathfrak{B} \times \mathfrak{D}, \mu \times P)$, by Theorem 7.17, restricting W to C, we have that $h \in L^2(Y \times \Omega, \mathfrak{C} \times \mathfrak{D}, L(\lambda \times \nu))$ and;

$\int_{Y \times \Omega} h^2(t,\omega) dL(\lambda \times \nu) = \int_{[0,1] \times \Omega} W^2(t,\omega) d\mu dP$

$= \int_{[0,1]} E(W_t^2) dt = \int_{[0,1]} t dt = \frac{1}{2}$ $(*)$

Moreover, using *-independence of the random variables $\{\omega_i\}_{1\leq i\leq\eta}$;

$\int_{Y\times\Omega} f^2(t,\omega)d(\lambda\times\nu) = \int_Y \frac{1}{\eta}\int_\Omega(\sum_{i=1}^{[\eta t]}\omega_i^2 + 2\sum_{1\leq i<j\leq[\eta t]}\omega_i\omega_j)d\nu d\lambda$

$= \int_Y \frac{[\eta t]}{\eta}d\lambda$

$= *\sum_{0\leq i<\eta}\frac{1}{\eta^2}[\eta(\frac{i}{\eta})]$

$= {}^*\sum_{0\leq i<\eta}\frac{i}{\eta^2} = \frac{1}{\eta^2}\frac{\eta-1}{2}(\eta-1) = \frac{1}{2}(\frac{\eta-1}{\eta})^2 \simeq \frac{1}{2}$ $(**)$

As $^\circ f = h$ a.e $L(\lambda\times\nu)$, using $(*)$,$(**)$, and Theorem 3.34(i), we have shown $(i)'$ as required.

Now applying Definition 7.20, we have;

$\int_0^1 W(t,\omega)dW(t,\omega) = {}^\circ\int_0^1 f(t,\omega)d\chi(t,\omega)$

$= {}^\circ(\sum_{i=0}^{\eta-1}\chi(\frac{i}{\eta},\omega)\frac{\omega_{i+1}}{\sqrt{\eta}})$

$= {}^\circ(\sum_{i=0}^{\eta-1}\sum_{j=1}^{i}\frac{\omega_j\omega_{i+1}}{\eta})$

$= {}^\circ\sum_{1\leq j<i\leq\eta-1}\frac{\omega_j\omega_i}{\eta}$

$= \frac{1}{2}{}^\circ(\sum_{1\leq j,i\leq\eta-1}\frac{\omega_j\omega_i}{\eta} - \sum_{j=1}^{\eta}\frac{\omega_j^2}{\eta})$

$= \frac{1}{2}{}^\circ((\sum_{j=1}^{\eta}\frac{\omega_j}{\sqrt{\eta}})^2 - \sum_{j=1}^{\eta}\frac{1}{\eta})$

$= \frac{1}{2}{}^\circ(\chi^2(1,\omega)-1) = \frac{1}{2}(W^2(1,\omega)-1)$

which agrees with the standard stochastic result.

Theorem 7.24. *Let $g\in\mathcal{F}$ and f a 2-lifting of g. Then, a.e P;*

$I(t,\omega) = \int_0^t g(\tau,\omega)dW(\tau,\omega)$

considered as a function on $[0,1]$ is continuous, where I is given in Definition 7.20. In fact, a.e P;

$F(t,\omega) = \int_0^t f(\tau,\omega)d\chi(\tau,\omega)$

considered as a function on $^*[0,1]$ is near standard in $^*C[0,1]$.

Proof. The result is analogous to Theorem 7.9. The proof relies on Doob's inequality for positive submartingales. □

Theorem 7.25. *Suppose $g \in \mathcal{F}$, and f is a 2-lifting of g, then $I \in \mathcal{F}$, I is $\mathfrak{B} \times \mathfrak{D}$, measurable ([43]), and ${}^\circ F$ is $L(\mathfrak{C}) \times \mathfrak{D}$-measurable, where F is given as in the proof of Theorem 7.22.*

Proof. We refer the reader to [1], part of the proof is required in the following Theorem 7.26.

□

One of the advantages of the non-standard approach to stochastic calculus, is that it allows one to show easily the following;

Theorem 7.26. *If $g \in \mathcal{G}_0$, and f is a 2-lifting of g, then $I(t, \omega)$, as in Definition 7.20, is equivalent, as a stochastic process, to a martingale, with respect to the filtration $\mathfrak{D}_t$, ([44]).*

Proof. Let I' be the modification of I, as given in the proof of Theorem 7.25. Then I' and agree I on $[0, 1] \times C$, where $P(C) = 1$, so they are equivalent as stochastic processes. We show that I' is a martingale.

(i) follows from the fact that I is $\mathfrak{B} \times \mathfrak{D}$ measurable, and $I = I'$ a.e $\mu \times L(\nu)$, $(*)$. Here, completeness of the product is required.

(ii). By the construction in the proof of Theorem 7.25, I'_t is measurable with respect to $\mathfrak{D}'_t \subset \mathfrak{D}_t$.

(iii). We have, for $t \in [0, 1]$;

$\int_\Omega I'^2(t, \omega) dL(\nu) = \int_\Omega I^2(t, \omega) dL(\nu)$

[43] We mean the complete product here.

[44] By which I mean a function $I : [0, 1] \times \Omega \rightarrow \mathcal{R}$, such that;

(i). I is $\mathfrak{B} \times \mathfrak{D}$ measurable (complete product).

(ii). I_t is measurable with respect to $\mathfrak{D}_t$, for $t \in [0, 1]$.

(iii). $E(|I_t|) < \infty$, for $t \in [0, 1]$.

(iv). $E(I_t | \mathfrak{D}_s) = I_s$, if $s < t$ belong to $[0, 1]$.

Most of this definition can be found in [39], see also [44] for a thorough discussion of discrete time martingales.

$= \int_\Omega {}^\circ F^2(t,\omega)d\nu$

$\leq {}^\circ \int_\Omega F^2(t,\omega)d\nu$

$= {}^\circ \int_\Omega \int_0^t f^2(t,\omega)d\lambda d\nu = ||g||^2_{L^2([0,t]\times\Omega)}$ (†)

using (∗), Definition 7.20, (see notation in Theorem 7.24), Theorem 3.16 and the proof of Theorem 7.22. Hence $I'_t \in L^2(\Omega, \mathfrak{D}, P)$, so $I'_t \in L^1(\Omega, \mathfrak{D}, P)$, by Holder's inequality, see [37].

(*iv*). Suppose $s < t$. We first show that $E(I'_t|\mathfrak{D}'_s) = I'_s$, (††). Suppose $i \in {}^*\mathcal{N}$, with $\frac{i}{\eta} \simeq s$, then we claim that $E(I'_t|\sigma(\mathfrak{A}_i)^{comp}) = I'_s$, (∗∗). As $I_t = I'_t$ a.e P, we have $E(I'_t|\sigma(\mathfrak{A}_i^{comp})) = E(I_t|\sigma(\mathfrak{A}_i^{comp}))$. We can also see that $F_t \in SL^2(\Omega, \mathfrak{A}, \nu)$. This follows from the calculation (†), Theorem 3.34(i), and the fact that;

$\int_\Omega I^2(t,\omega)dL(\nu) = ||g||^2_{L^2([0,t]\times\Omega)}$

by Ito's isometry, as $g \in \mathcal{G}_0$. Hence, by Theorem 3.34(iv), $F_t \in SL^1(\Omega, \mathfrak{A}, \nu)$, (∗∗∗). Applying Theorem 7.3(ii) and (∗∗∗);

$E(I_t|\sigma(\mathfrak{A}_i)^{comp}) = E({}^\circ F_t|\sigma(\mathfrak{A}_i)^{comp}) = {}^\circ E(F_t|\mathfrak{A}_i)$

We have;

$E(F_t|\mathfrak{A}_i) = \sum_{j=0}^{i-1} f(\frac{j}{\eta},\omega)\frac{\omega_{j+1}}{\sqrt{\eta}}$

by *-independence of the sequence $\{\omega_j\}_{0\leq j\leq[\eta t]+1}$. Letting $s' = \frac{i-1}{\eta}$, so $s' \simeq s$, $E(F_t|\mathfrak{A}_i) = F_{s'}$. We have, using Theorem 7.24, that $I_s = I_{s'}$ a.e P, so $I'_s = I_s = I_{s'}$ a.e P. As I'_s is $\sigma(\mathfrak{A}_i)^{comp}$-measurable, we have $E(I'_t|\mathfrak{A}_i^{comp}) = I'_s$, showing (∗∗). As $\mathfrak{D}'_s \subset \sigma(\mathfrak{A}_i)^{comp}$, and I'_s is $\mathfrak{D}'_s$-measurable, we have $E(I'_t|\mathfrak{D}'_s) = I'_s$, showing (††).

If $A \in \mathfrak{D}_s$, then, by Lemma 7.15(i), $A \in \mathfrak{D}'_{s_1}$, for $s < s_1 < t$. As $E(I'_t|\mathfrak{D}'_{s_1}) = I'_{s_1}$, to show (*iv*), it is sufficient to prove that;

$\int_A I'_s dL(\nu) = lim_{s_1\rightarrow s} \int_A I'_{s_1} dL(\nu)$ (†††)

To show (†††), observe that $||I'_{s_1} - I'_s||_2^2 \leq ||g_{[0,s_1]} - g_{[0,s]}||_2^2$ by (†), where $g_{[0,s_1]}$ is obtained by truncating the function g to the interval $[0, s_1]$, ([45]). Using Holder's inequality and the DCT, we have $lim_{s_1 \rightarrow s}||I'_{s_1} - I'_s||_1 \leq lim_{s_1 \rightarrow s}||g_{[0,s_1]} - g_{[0,s]}||_1 = 0$. Therefore, (†††) is shown. This proves (iv).

□

We finally prove a generalisation of Ito's Theorem. This covers the case not shown in the paper [1].

Theorem 7.27. *Ito's Theorem*

Suppose $h : \mathcal{R}^r \rightarrow \mathcal{R}$ belongs to C^2, with continuous partial derivatives $\{h_k, h_{kl}\}_{1 \leq k,l \leq r}$, W is the Brownian motion given in Definition 7.14, $\{g_i\}_{1 \leq i \leq r}$ belong to $\mathcal{G}$, and $\{a_i\}_{1 \leq i \leq r}$ belong to $L^1([0,1] \times \Omega, \mathfrak{B} \times \mathfrak{D}, \mu \times P)$, with $a_{i,t}$ $\mathfrak{D}_t$-measurable. Let;

$I_i(t, \omega) = \int_0^t a_i(\tau, \omega) d\tau + \int_0^t g_i(\tau, \omega) dW(\tau, \omega)$, $(1 \leq i \leq r)$.

$I(t, \omega) = (I_1(t, \omega), \ldots, I_r(t, \omega))$

$H(t, \omega) = h(I(t, \omega))$

Then H is a stochastic integral;

$H(t, \omega) - H(0, \omega)$

$= \sum_{k=1}^r \int_0^t h_k(I(\tau, \omega)) a_k(\tau, \omega) d\tau$

$+ \sum_{k=1}^r \int_0^t h_k(I(\tau, \omega)) g_k(\tau, \omega) dW(\tau, \omega)$

$+ \frac{1}{2} \sum_{1 \leq k,l \leq r} \int_0^t h_{kl}(I(\tau, \omega)) g_k g_l(\tau, \omega) d\tau$

Proof. We assume $r = 2$, the general case follows the same method. Let $\{f_i\}_{1 \leq i \leq r}$ and $\{b_i\}_{1 \leq i \leq r}$ be liftings of $\{g_i\}_{1 \leq i \leq r}$ and $\{a_i\}_{1 \leq i \leq r}$ respectively, as guaranteed by Theorem 7.19. Let $\{F_i\}_{1 \leq i \leq r}$ be as in

[45] Technically, you need to show that I_{s_1} is the non standard stochastic integral of $g_{[0,s_1]}$, and then apply Theorem 7.22, however, this is clear by truncating the corresponding lift of g.

Remarks 7.21. Let;

$D_i(t,\omega)=\int_0^t b_i(\tau,\omega)d\lambda(\tau)$, ([46])

$G_i(t,\omega)=D_i(t,\omega)+F_i(t,\omega)$ $(1\leq i\leq r)$

Then a.e P, for all $t\in {}^*[0,1]$, ${}^\circ G_i(t,\omega)=I_i({}^\circ t,\omega)$, $(\dagger)$, ([47]).

Using Lemma 2.12(i) and Theorem 2.25 (footnote 4), there exists $M\in {}^*\mathcal{N}\setminus\mathcal{N}$ such that;

$$\epsilon=sup_{(||\bar{t}||_r<M,||\bar{\delta}||_r<2\sqrt{r}\eta^{-\frac{1}{3}})}max_{1\leq k,l\leq r}|{}^*h_{kl}(\bar{t}+\bar{\delta})-{}^*h_{kl}(\bar{t})|\simeq 0\ (\sharp)$$

[46]Technically the internal integral is not defined when $[\eta t]\neq\eta t$, however, one can easily extend the definition by;

$\int_0^t b_i(\tau,\omega)d\lambda(\tau)=\int_0^{\frac{[\eta t]}{\eta}} b_i(\tau,\omega)d\lambda(\tau)+\frac{\eta t-[\eta t]}{\eta}b_i(\frac{[\eta t]}{\eta},\omega)$

$=\frac{1}{\eta}\sum_{j=0}^{[\eta t]-1} b_i(\frac{j}{\eta},\omega)+\frac{\eta t-[\eta t]}{\eta}b_i(\frac{[\eta t]}{\eta},\omega)$

[47]The proof of $(\dagger)$ is similar to parts of the proof of (AA) and (B) below. We have to show, first, that;

${}^\circ D_i(t,\omega)=\int_0^{\circ t} a_i(\tau,\omega)d\tau$. (a.e P, for all $t\in {}^*[0,1]$). $(*)$

By the construction of b_i in Theorem 7.19, we have that;

${}^\circ b_i(\tau,\omega)=a_i'(\tau,\omega)=a_i({}^\circ\tau,\omega)$ a.e $L(\lambda\times\nu)$

Using [37], Theorem 8.8(c), and Theorem 7.17, we have ,a.e $P(\omega)$, that $a'_{i,\omega}\in L^1(Y,\mathfrak{C},\mu)$ and, for such ω, ${}^\circ b_{i,\omega}=a'_{i,\omega}$, a.e $L(\lambda)$. Letting $\phi_i(\omega)=\int_Y b_i(\tau,\omega)d\lambda$ and $\psi_i(\omega)=\int_Y a_i'(\tau,\omega)dL(\lambda)$, we can use the fact that $b_i\in SL^1(Y\times\Omega,\mathfrak{C}\times\mathfrak{D},\lambda\times\nu)$, a_i' is $\mathfrak{C}\times\mathfrak{D}$-measurable (complete product), [37] Theorem 8.8(a) and Theorem 8.12 Lemma 1 (Fubini's Theorem), to show that for all $A\in\mathfrak{A}$, ${}^\circ\int_A\phi_i dL(\nu)=\int_A\psi_i dL(\nu)$. Hence, by Lemma 3.37, ${}^\circ\phi_i=\psi_i$ a.e $L(\lambda)$. $(*)$ now follows easily from Theorem 3.34(i), then Theorem 3.24(ii)', and, finally, Theorem 7.17.

We also have to show that;

${}^\circ F_i(t,\omega)=\int_0^{\circ t} g_i(\tau,\omega)dW(\tau,\omega)$. (a.e P, for all $t\in {}^*[0,1]$). $(**)$

This is easier, using Theorem 7.24, we have a.e $P(\omega)$, that $F_i(t,\omega)$ is near standard. Hence, for such ω, ${}^\circ F_i(t,\omega)={}^\circ F_i({}^\circ t,\omega)$. Now the result $(**)$ follows just from Definition 7.20.

Using Theorem 3.30(i), we can assume that;

$max_{1\leq i\leq r}\{|f_i|,|b_i|\}<min\{\eta^{\frac{1}{6}},\epsilon^{-\frac{1}{3}}\}$

Then, for $1\leq i\leq r$;

$|F_i(\frac{j}{\eta},\omega)-F_i(\frac{j-1}{\eta},\omega)|=|f_i(\frac{j-1}{\eta},\omega)\frac{\omega_j}{\eta}|<min\{\eta^{-\frac{1}{3}},\epsilon^{-\frac{1}{3}}\eta^{\frac{-1}{2}}\}$

$|D_i(\frac{j}{\eta},\omega)-D_i(\frac{j-1}{\eta},\omega)|=|\int_{\frac{j-1}{\eta}}^{\frac{j}{\eta}}b_i(\tau,\omega)d\lambda(\tau)|$

$<\frac{1}{\eta}|b_i(\frac{j-1}{\eta},\omega)|<min\{\eta^{-\frac{1}{3}},\epsilon^{-\frac{1}{3}}\eta^{\frac{-1}{2}}\}$

$|G_i(\frac{j}{\eta},\omega)-G_i(\frac{j-1}{\eta},\omega)|<min\{2\eta^{-\frac{1}{3}},2\epsilon^{-\frac{1}{3}}\eta^{\frac{-1}{2}}\}$

so $max_{1\leq i\leq r}\{|G_i(\frac{j}{\eta},\omega)-G_i(\frac{j-1}{\eta},\omega)|\}<min\{2\eta^{-\frac{1}{3}},2\epsilon^{-\frac{1}{3}}\eta^{\frac{-1}{2}}\}$ ($\sharp\sharp$)

Let;

$\bar{G}(t,\omega)=(G_1(t,\omega),\ldots,G_r(t,\omega))$

$M_\omega=sup_{t\in{}^*[0,1]}max_{1\leq i\leq r,1\leq k,l\leq r}\{|G_i|,|{}^*h(\bar{G})|,|*h_k(\bar{G})|,|{}^*h_{kl}(\bar{G})\}$

Using (†), Theorem 7.24, and the fact that $\int_0^t a_i(\tau,\omega)d\tau$ is continuous as a function of t, (for any given ω), we have, a.e P, that G_i is near standard in ${}^*C[0,1]$. By Theorem 2.25(footnote 4), it follows easily that if each G_i is near standard, then ${}^*h(\bar{G})$ is near standard, and similarly, for the higher order terms. Hence, we have a.e P, that M_ω is finite, ($\sharp\sharp\sharp$). Then a.e P;

$H(t,\omega)-H(0,\omega)$

$={}^*h(I(t,\omega))-{}^*h(I(0,\omega))$

$={}^*h({}^\circ\bar{G}(t,\omega))-{}^*h({}^\circ\bar{G}(0,\omega))$

$={}^\circ({}^*h(\bar{G}(t,\omega))-{}^*h(\bar{G}(0,\omega)))$

$={}^\circ({}^*h(\bar{G}(\frac{[\eta t]}{\eta},\omega))-{}^*h(\bar{G}(0,\omega)))$ $(*)$

$={}^\circ((\sum_{j=1}^{[\eta t]}{}^*h(\bar{G}(\frac{j}{\eta},\omega))-{}^*h(\bar{G}(\frac{j-1}{\eta},\omega))))$ $(**)$

$= {}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r}{}^*h_k(\bar{G}(\frac{j-1}{\eta},\omega))(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega)))))$

$+{}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r}(\frac{1}{2}{}^*h_{kk}(\bar{G}(\frac{j-1}{\eta},\omega))+\epsilon_k)(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega))^2)))$

$+{}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{1\leq k<l\leq r}({}^*h_{kl}(\bar{G}(\frac{j-1}{\eta},\omega))+\epsilon_{kl})(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega))$

$(G_l(\frac{j}{\eta},\omega)-G_l(\frac{j-1}{\eta},\omega))))) \ (***)$

In $(*)$, we have used the fact that ${}^*h(\bar{G})(t,\omega)$ is near standard in ${}^*C[0,1]$, so ${}^*h(\bar{G})(\frac{[\eta t]}{\eta},\omega)\simeq {}^*h(\bar{G})(t,\omega)$. $(**)$ uses the same trick as in the usual approach, see [39], to write the expression $(*)$ as an alternating sum. Finally, $(***)$ uses the multivariate Taylor's formula, transferred to ${}^*C^2({}^*\mathcal{R}^r,{}^*\mathcal{R})$, where $\{\epsilon_k,\epsilon_{kl}\}$ denote error terms, ([48]).

We first show that we can ignore the error terms $\{\epsilon_k,\epsilon_{kl}\}$ in $(***)$ Using footnote 41, (♯),(♯♯), (♯♯♯), we have;

---

[48]The multivariate Taylor's formula, see [43], says that if h is C^2, in r variables, then;

$h(\bar{u})=h(\bar{s})+\sum_{k=1}^{r}h_k(\bar{s})(u_k-s_k)+\sum_{|\alpha|=2}R_\alpha(\bar{u})(\bar{u}-\bar{s})^\alpha$

where α is a multi-index and the remainder term is given by;

$R_\alpha(\bar{u})=\frac{|\alpha|}{\alpha!}\int_0^1(1-w)^{|\alpha|-1}D^\alpha h(\bar{s}+w(\bar{u}-\bar{s}))dw$

We can split the above sum into terms α_k, $1\leq k\leq r$, involving repeated indices, and α_{kl}, $1\leq k<l\leq r$. Then;

$R_{\alpha_k}(\bar{u})(u_k-s_k)^2=\frac{1}{2}h_{kk}(\bar{s})(u_k-s_k)^2+(R_{\alpha_k}(\bar{u})-\frac{1}{2}h_{kk}(\bar{s}))(u_k-s_k)^2$

$R_{\alpha_{kl}}(\bar{u})(u_k-s_k)(u_l-s_l)$

$=h_{kl}(\bar{s})(u_k-s_k)(u_l-s_l)+(R_{\alpha_{kl}}(\bar{u})-h_{kl}(\bar{s}))(u_k-s_k)(u_l-s_l)$

Letting $\epsilon_k=R_{\alpha_k}(\bar{u})-\frac{1}{2}h_{kk}(\bar{s})$, we have that;

$|\epsilon_k|=|\int_0^1(1-w)(h_{kk}(\bar{s}+w(\bar{u}-\bar{s}))-h_{kk}(\bar{s}))dw|$

$\leq sup_{(||\bar{\delta}||_r\leq||\bar{u}-\bar{s}||_r)}|h_{kk}(\bar{s}+\bar{\delta})-h_{kk}(\bar{s})|$

where we have used the fact that $\int_0^1(1-w)dw=\frac{1}{2}$. Similarly, letting $\epsilon_{kl}=R_{\alpha_{kl}}(\bar{u})-h_{kl}(\bar{s})$, we have;

$|\epsilon_{kl}|\leq 2sup_{(||\bar{\delta}||_r\leq||\bar{u}-\bar{s}||_r)}|h_{kl}(\bar{s}+\bar{\delta})-h_{kl}(\bar{s})|$

$\sum_{j=1}^{[\eta t]}|\epsilon_k|(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega))^2$

$\leq \eta(4\epsilon^{\frac{-2}{3}}\eta^{-1})sup_{(||\bar{t}||_r<M,||\bar{\delta}||_r<||2\bar{\eta}^{\frac{-1}{3}}||)}|^*h_{kk}(\bar{t}+\bar{\delta})-{}^*h_{kk}(\bar{t})|$

$=4\epsilon^{\frac{-2}{3}}sup_{||\bar{\delta}||_r<2\sqrt{r}\eta^{\frac{-1}{3}}}|^*h_{kk}(\bar{t}+\bar{\delta})-{}^*h_{kk}(\bar{t})|=4\epsilon^{\frac{-2}{3}}\epsilon=4\epsilon^{\frac{1}{3}}\simeq 0$

The calculation for ϵ_{kl} is the same (except with an extra factor of 2). We now evaluate the remaining three remaining sums in $(***)$.

(A). $^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r}{}^*h_k(\bar{G}(\frac{j-1}{\eta},\omega))(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega)))))$

$=^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r}{}^*h_k(\bar{G}(\frac{j-1}{\eta},\omega))(\frac{b_k(\frac{j-1}{\eta},\omega)}{\eta}))))$, (AA)

$+^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r}{}^*h_k(\bar{G}(\frac{j-1}{\eta},\omega))(f_k(\frac{j-1}{\eta},\omega)\frac{\omega_j}{\sqrt{\eta}})))))$ (AB)

To obtain these terms, we have just used the definition of G_k, Remarks 7.21, and footnote 46.

(AA). For $n\in\mathcal{N}$, let $A_n=\{\omega: M_\omega<n\}$. Then A_n is internal and $P(\bigcup_{n=1}^{\infty}A_n)=1$. If $B\in\mathfrak{A}$, let $B_n=A_n\cap B$. By the definition of M_ω, the fact that $b_k(\tau,\omega)\in SL^1(Y\times\Omega,\mathfrak{C}\times\mathfrak{A},\lambda\times\nu)$, and Theorem 3.30(i), we have that $b_k.(^*h_k\circ\bar{G})|_{Y\times A_n}\in SL^1(Y\times A_n,\mathfrak{C}\times\mathfrak{A},\lambda\times\nu)$. Let $\theta_k(\tau,\omega)={}^*h_k(\bar{G}(\frac{[\eta\tau]}{\eta},\omega))b_k(\frac{[\eta\tau]}{\eta},\omega)$. Notice that, a.e in A_n, P;

$^\circ\theta_k(\tau,\omega)={}^{\circ*}h_k(\bar{G}(\frac{[\eta\tau]}{\eta},\omega))^\circ b_k(\frac{[\eta\tau]}{\eta},\omega)=h_k(I(^\circ\tau,\omega))a_k(^\circ\tau,\omega)$ $(\dagger\dagger)$

by $(\dagger)$, the fact that b_k is a lift of a_k ((ii)',Definition 7.18), and continuity of h, see Theorem 2.25. As we have already remarked, $\theta_k|_{Y\times A_n}\in SL^1(Y\times A_n,\mathfrak{C}\times\mathfrak{A},\lambda\times\nu)$.

Moreover, for $\omega\in A_n$, using footnote 46;

$\phi_k(t,\omega)=\int_0^t\theta_k(\tau,\omega)d\lambda=\sum_{j=1}^{[\eta t]}h_k(\bar{G}(\frac{j-1}{\eta},\omega))(\frac{b_k(\frac{j-1}{\eta},\omega)}{\eta})$ $(****)$

and;

$^\circ\int_{B_n}\phi_k(t,\omega)d\nu$

$=^\circ\int_{[0,t]\times B_n}\theta_k(\tau,\omega)d(\lambda\times\nu)$ (definition of ϕ_k)

$= \int_{[0,t]\times B_n} {}^\circ\theta_k(\tau,\omega)dL(\lambda\times\nu)$ $(\theta_k \in SL^1(Y\times A_n, \mathfrak{C}\times\mathfrak{A}, \lambda\times\nu))$

$= \int_{[0,t]\times B_n} {}^\circ\theta_k(\tau,\omega)dL(\lambda)\times L(\nu)$ $(\dagger\dagger\dagger)$

$= \int_{B_n}\int_0^t h_k(I(\tau,\omega))a_k(\tau,\omega)d\tau dP$, (by $(\dagger\dagger)$ and Theorem 7.17)

$= \int_{B_n}\psi_k(t,\omega)dP$

where $\psi_k(t,\omega) = \int_0^t h_k(I(\tau,\omega))a_k(\tau,\omega)d\tau$. In making the step to $(\dagger\dagger\dagger)$, we have used Theorem 7.25, and the fact that ${}^\circ b_k$ is $L(\lambda)\times L(\nu)$-measurable, see the construction in Theorem 7.19. By Lemma 3.37, ${}^\circ\phi_k = \psi_k$ a.e A_n, therefore a.e P. By, $(****)$, we obtain;

$${}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r} {}^*h_k(\bar{G}(\tfrac{j-1}{\eta},\omega))(\tfrac{b_k(\frac{j-1}{\eta},\omega)}{\eta}))))$$

$$= \sum_{k=1}^{r}\int_0^t h_k(I(\tau,\omega))a_k(\tau,\omega)d\tau$$

(AB). We have that $F_k(\frac{[\eta\tau]}{\eta},\omega)|_{[0,s]}$ is $\mathfrak{C}\times\mathfrak{A}_{[\eta s]}$-measurable, for all $s\in {}^*[0,1]$, as $f_k(\tau,\omega)$ satisfies this condition, and we can then use [37], Theorem 8.8(a). The same argument applies for $D_k(\frac{[\eta\tau]}{\eta},\omega)$, using the corresponding property of $b_k(\tau,\omega)$. It follows, using the same argument as in (AA), that ${}^*h_k(\bar{G}(\frac{[\eta\tau]}{\eta},\omega))f_k(\frac{[\eta\tau]}{\eta},\omega)|_{Y\times A_n}$ is a 2-lifting of $h_k(I(\tau,\omega))g_k(\tau,\omega)|_{[0,1]\times A_n}$, in the sense of Definition 7.18. Now applying Remarks 7.21 and Definition 7.20, we have that;

$${}^\circ((\sum_{j=1}^{[\eta t]} {}^*h_k(\bar{G}(\tfrac{j-1}{\eta},\omega))(f_k(\tfrac{j-1}{\eta},\omega)\tfrac{\omega_j}{\sqrt{\eta}})))$$

$$= {}^\circ\int_0^t {}^*h_k(\bar{G}(\tfrac{[\eta\tau]}{\eta},\omega))f_k(\tfrac{[\eta\tau]}{\eta},\omega)d\chi(\tau,\omega)$$

$$= \int_0^t h_k(I(\tau,\omega))g_k(\tau,\omega)dW(\tau,\omega)$$

a.e A_n, P, hence a.e Ω, P. Therefore;

$${}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r} {}^*h_k(\bar{G}(\tfrac{j-1}{\eta},\omega))(f_k(\tfrac{j-1}{\eta},\omega)\tfrac{\omega_j}{\sqrt{\eta}}))))$$

$$= \sum_{k=1}^{r}\int_0^t h_k(I(\tau,\omega))g_k(\tau,\omega)dW(\tau,\omega)$$

(B) ${}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^{r}(\frac{1}{2}{}^*h_{kk}(\bar{G}(\frac{j-1}{\eta},\omega)))(G_k(\frac{j}{\eta},\omega) - G_k(\frac{j-1}{\eta},\omega))^2)))$

We expand, using the definition of G_k, footnote 46, and Remarks 7.21 ;

$(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega))^2$

$=(\frac{b_k(\frac{j-1}{\eta},\omega)}{\eta}+f_k(\frac{j-1}{\eta},\omega)\frac{\omega_j}{\sqrt{\eta}})^2$

$=\frac{b_k^2(\frac{j-1}{\eta},\omega)}{\eta^2}+\frac{f_k^2(\frac{j-1}{\eta},\omega)}{\eta}+\frac{2b_kf_k(\frac{j-1}{\eta},\omega)\omega_j}{\eta\sqrt{\eta}}$

We have, a.e P, and using the bounds on $\{|b_k|,|f_k|\}$;

$|\sum_{j=1}^{[\eta t]}\frac{1}{2}{}^*h_{kk}(\bar{G}(\frac{j-1}{\eta},\omega))\frac{b_k^2(\frac{j-1}{\eta},\omega)}{\eta^2}|\leq\eta M_\omega\frac{(\eta^{\frac{1}{6}})^2}{2\eta^2}=\frac{M_\omega}{2\eta^{\frac{2}{3}}}\simeq 0$

$|\sum_{j=1}^{[\eta t]}\frac{1}{2}{}^*h_{kk}(\bar{G}(\frac{j-1}{\eta},\omega))\frac{2b_kf_k(\frac{j-1}{\eta},\omega)\omega_j}{\eta\sqrt{\eta}}|\leq\eta M_\omega\frac{(\eta^{\frac{1}{6}})^2}{\eta^{\frac{3}{2}}}=\frac{M_\omega}{\eta^{\frac{1}{6}}}\simeq 0$

So we are left with;

${}^\circ(\sum_{j=1}^{[\eta t]}\frac{1}{2}{}^*h_{kk}(\bar{G}(\frac{j-1}{\eta},\omega))\frac{f_k^2(\frac{j-1}{\eta},\omega)}{\eta})$

$=\frac{1}{2}\int_0^t h_{kk}(I(\tau,\omega))g_k^2(\tau,\omega)d\tau$

exactly, as in the argument of (AA). This gives;

${}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{k=1}^r(\frac{1}{2}{}^*h_{kk}(\bar{G}(\frac{j-1}{\eta},\omega)))(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega))^2)))$

$=\frac{1}{2}\sum_{k=1}^r\int_0^t h_{kk}(I(\tau,\omega))g_k^2(\tau,\omega)d\tau$

(C) ${}^\circ((\sum_{j=1}^{[\eta t]}(\sum_{1\leq k<l\leq r}({}^*h_{kl}(\bar{G}(\frac{j-1}{\eta},\omega)))(G_k(\frac{j}{\eta},\omega)-G_k(\frac{j-1}{\eta},\omega))$

$(G_l(\frac{j}{\eta},\omega)-G_l(\frac{j-1}{\eta},\omega)))))$

The calculation is almost identical to (B), the terms involving $\{b_kb_l,b_kf_l\}_{k<l}$ vanish and we are left with the terms involving $\{f_kf_l\}_{k<l}$. We then add in the factor $\frac{1}{2}$ by summing over all $k\neq l$. This gives the result, ([49]). □

[49]If we consider the case $r=2$ and take $a_1=1$,$a_2=0$, $g_1=0$, $g_2=1$, we obtain $I_1(t,\omega)=\int_0^t d\tau=t$ and $I_2(t,\omega)=\int_0^t dW(\tau,\omega)=W_t$. Then, the theorem gives us;

If $X_t=h(t,W_t)$;

Remarks 7.28. *One could hope to use some of the ideas presented in this section, in the context of fractional Brownian motion. This has already been initiated in* [13]. *Another potentially related application is that of finding a continuous time analogue of the GARCH model in financial econometrics, see* [14]. *As far as I am aware, there has been no work in nonstandard analysis, done in this direction.*

References

[1] A Non-Standard Representation for Brownian Motion and Ito Integration, Robert Anderson, Isreal Journal of Mathematics, (1976).

[2] StarFinite Representations of Measure Spaces, Robert Anderson, Transactions of the American Mathematical Society, (1982).

[3] Introduction to Real Analysis, Robert Bartle and Donald Schubert, John Wiley, (1982).

[4] Methods for Stochastics and Finance, Mitchell Berger, available on ELE, (2012).

[5] The Analyst; or a Discourse Addressed to an Infidel Mathematician, George Berkeley, (1734).

[6] Mathematical Theory of Option Pricing, Nigel Byott, available on ELE, (2011).

[7] A Non-Standard Integration Theory for Unbounded Functions, Allen Bernstein, Zeitschrift fur mathematische Logik und Grundlagen der Mathematik, (1974).

[8] Nonstandard Measure Theory, Allen Bernstein and Frank Wattenberg, Applications of Model Theory to Algebra, Probability Theory and Analysis,

---

$X_t - X_0 = \int_0^t \frac{\partial h}{\partial \tau} + \frac{1}{2}\frac{\partial^2 h}{\partial {W_\tau}^2} d\tau + \int_0^t \frac{\partial h}{\partial W_\tau} dW_\tau$

Good accounts of the applications of this result, including the Black Scholes formula for option pricing, are given in [17] and [10]. In [1], the case when $r = 1$, but with two independent Brownian motions is considered. The idea is to show that the cross terms involving $\{g^1, g^2\}$, corresponding to $\{W_1, W_2\}$, can be written as $\phi(\frac{j-1}{\eta}, \omega)\frac{\omega_{1,j}\omega_{2,j}}{\sqrt{\eta}}$, where $\phi(\frac{j-1}{\eta}, \omega) \simeq 0$. It is easy to show by *-independence of the sequence $\{\omega_{1,j}, \omega_{2,j}\}$, that the products $\frac{\omega_{1,j}\omega_{2,j}}{\sqrt{\eta}}$ correspond to the increments of a new random walk χ'. Checking that $\phi(\frac{[\eta\tau]}{\eta}, \omega)$ does define a lifting of 0 in the sense of Definition 7.18, we write the *-finite sum as a stochastic integral $\int_0^t \phi(\frac{[\eta\tau]}{\eta}, \omega) d\chi'(\tau, \omega)$. By Theorem 7.22, we then know that the standard part of this sum is 0 as well. Hence, we can ignore these terms. The rest of the calculation is similar to the above proof of the theorem.

(1968).

[9] Linear Analysis, Bela Bollobas, Cambridge Mathematical Textbooks, (1990).

[10] Mathematical Theory of Option Pricing, Nigel Byott, available on ELE, (2011).

[11] Model Theory, Chen Chang and Jerome Keisler, Studies in Logic and the Foundation of Mathematics, North Holland, (1977).

[12] Loeb Measures in Practice: Recent Advances, Nigel Cutland, Lecture Notes in Mathematics, Springer, (2001).

[13] Stock Price Returns and the Joseph Effect, a fractional version of the Black Scholes model, Nigel Cutland, Ekkehard Kopp, Walter Willinger, Progress in Probability 36(1995), 327-351.

[14] Lecture Notes on Advanced Econometrics, James Davidson, available on ELE, (2013).

[15] Probability, An Introduction, Geoffrey Grimmett, Domini Welsh, Oxford Science Publications, (1986).

[16] Logic for Mathematicians, A.G.Hamilton, Cambridge University Press, (1988).

[17] Lecture Notes in Advanced Probability and Stochastic Calculus, Mark Holland, available on ELE, (2013)

[18] The Mordell Lang Conjecture for Function Fields, Ehud Hrushovski, Journal of the American Mathematical Society, (1996).

[19] The Manin Mumford Conjecture and the Model Theory of Difference Fields, Ehud Hrushovski, Annals of Pure and Applied Logic, (2001).

[20] Zariski Geometries, Ehud Hrushovski and Boris Zilber, Journal of The American Mathematical Society, (1996).

[21] A Simple Proof of the Ergodic Theorem Using Nonstandard Analysis, Teturo Kamae, Isreal Journal of Mathematics, (1982).

[22] An Infinitesimal Approach to Stochastic Analysis, Jerome Keisler, Memoirs of the American Mathematical Society, (1984).

[23] Fourier Analysis, Tom Korner, CUP, (1998).

[24] Conversion from Nonstandard to Standard Measure Spaces and Applications in Probability Theory, Peter Loeb, Transactions of the American Mathematical Society, (1975).

[25] A Nonstandard Analysis Approach to Fourier Analysis, Wilhelm Luxemburg, Studies in Logic and The Foundations of Mathematics, (1972).

[26] The Quadrature of Curves, Isaac Newton, (1704).

[27] Analysis of Equations of an Infinite Number of Terms, Isaac Newton, written in 1669, a published version appearing in Latin in 1711.

[28] The Calculus Becomes an Algorithm, Isaac Newton, (middle 1665?).

[29] The General Problem of Tangents, Curvature and Limit-Motion Analysed by the Method of Fluxions, Isaac Newton, (October 1665-May 1666).

[30] Markov Chains, James Norris, Cambridge Series in Statistical and Probabilistic Mathematics, (1997).

[31] Ergodic Theory, Karl Petersen, Cambridge Studies in Advanced Mathematics, (1983).

[32] A Non Standard Bezout Theorem for Curves, Tristram de Piro, AG/LO ArXiv(0406176), (2004).

[33] An Interpretation of Newton's Work in Calculus, Tristram de Piro, available on my website http://www.magneticstrix.net, (2008).

[34] A Nonstandard Approach to the Theory of Algebraic Curves, Tristram de Piro, available on request, submitted to the Journal of Logic and Analysis, (2013).

[35] Some Notes on Non-Standard Differentiation and Integration, Tristram de Piro, available on my website, (2008).

[36] Non-Standard Analysis, Abraham Robinson, Studies in Logic and the Foundation of Mathematics, North-Holland, (1974).

[37] Real and Complex Analysis, Walter Rudin, McGraw Hill Third Edition, (1987).

[38] Trattato di geometria algebrica, 1: Geometria delle serie lineari, Francesco Severi, Zanichelli, Bologna, (1926).

[39] Stochastic Calculus and Financial Applications, Michael Steele, Applications of Mathematics, Springer, (2001).

[40] Fourier Analysis, An Introduction, Elias Stein and Rami Shakarchi, Princeton Lectures in Analysis, (2003).

[41] Nonstandard Measure Theory: avoiding pathological sets, Frank Wattenberg, Transactions of the American Mathematical Society, (1979).

[42] Lebesgue Integration and Measure, Alan Weir, Cambridge University Press, (1973).

[43] Wikipedia; Taylor's Theorem.

[44] Probability with Martingales, David Williams, Cambridge Mathematical Textbooks, (1991).

[45] Zariski Geometries; Geometry from the Logician's Point of View, Boris Zilber, London Mathematical Society Lecture Note Series, (2010)

Mathematics Department, The University of Exeter, Exeter
E-mail address: tdpd201@exeter.ac.uk