1. Loeb Measures

1.1 Introduction

Loeb measures, discovered by Peter Loeb in 1975 [71], are very rich standard measure spaces, constructed using nonstandard analysis (NSA). The range of fields in which they have found significant applications is vast, including measure and probability theory, stochastic analysis, differential equations (ordinary, partial and stochastic), functional analysis, control theory, mathematical physics, economics and mathematical finance theory.

The richness of Loeb measures makes them good for

- *constructing* measures with special properties (for example the *rich probability spaces* of Fajardo & Keisler [49, 50, 62]);

- *representing* complex measures in ways that make them more manageable (for example Wiener measure) – see section 1.3.3 below;

- *modelling* physical and other phenomena;

- *proving existence results* in analysis – for example solving differential equations (DEs) of all kinds (including partial DEs, stochastic DEs and even stochastic partial DEs) and showing the existence of attractors.

Later lectures will describe some recent uses of Loeb spaces that illustrate these themes – in fluid mechanics, in stochastic calculus of variations and related topics, and in mathematical finance theory.

This lecture will outline the basic Loeb measure construction and give some simple applications, with a little of the theory of Loeb integration. From one point of view, Loeb measures are simply ultraproducts of previously given measure spaces, such as were considered in an early paper of Dacunha-Castelle & Krivine [46]. The rôle of NSA in their construction is to provide a systematic way to understand their properties, which opens the way for efficient and powerful applications; without this we would have a supply of very rich measure spaces but only *ad hoc* means to comprehend them.

Necessarily these lectures will be somewhat informal and lacking in a great deal of detail. The aim is to convey some of the basic ideas and flavour of Loeb measures and how they work, as well as pointing to the literature where the topics can be pursued in depth. We must begin with a brief and informal look at NSA itself.

N.J. Cutland: LNM 1751, pp. 1–28, 2000.

1.2 Nonstandard Analysis

1.2.1 The hyperreals

Nonstandard analysis (discovered by Abraham Robinson in 1960 [83]) begins with the construction of a richer real line ${}^*\mathbb{R}$ called the *hyperreals* or *nonstandard reals*. This is an ordered field that extends the (standard) reals $\mathbb{R}$ in two main ways:

(1) ${}^*\mathbb{R}$ contains non-zero *infinitesimal numbers*; and

(2) ${}^*\mathbb{R}$ contains positive and negative *infinite numbers*.

This is made precise by the following definitions (where $|\cdot|$ is the extension[1] of the modulus function to ${}^*\mathbb{R}$).

Definition 1.1 Let $x \in {}^*\mathbb{R}$. We say that

(i) x is *infinitesimal* if $|x| < \varepsilon$ for all $\varepsilon > 0$, $\varepsilon \in \mathbb{R}$;

(ii) x is *finite* if $|x| < r$ for some $r \in \mathbb{R}$;

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

(iv) We say that x and y are *infinitely close*, denoted by $x \approx y$, if $x - y$ is infinitesimal. So $x \approx 0$ is another way to say that x is infinitesimal.

(v) The *monad* of a real number r is the set

$$\text{monad}(r) = \{x : x \approx r\}$$

of hyperreals that are infinitely close to r. Thus monad(0) is the set of infinitesimals, and $\text{monad}(r) = r + \text{monad}(0)$.

Of course, once a field has non-zero infinitesimals, then there must be infinite elements also – these are the reciprocals of infinitesimals. It follows also that $\mathbb{R}$ is enriched in having, for each $r \in \mathbb{R}$, new elements x with $x \approx r$ (taking $x = r + \delta$ where δ is infinitesimal).

One way to construct ${}^*\mathbb{R}$ is as an *ultrapower* of the reals

$${}^*\mathbb{R} = \mathbb{R}^{\mathbb{N}}/\mathcal{U}$$

where $\mathcal{U}$ is a nonprincipal ultrafilter[2] (or maximal filter) on $\mathbb{N}$.

That is, ${}^*\mathbb{R}$ consists of equivalence classes of sequences of reals under the equivalence relation $\equiv_{\mathcal{U}}$, defined by

$$(a_n) \equiv_{\mathcal{U}} (b_n) \qquad \Longleftrightarrow \qquad \{n : a_n = b_n\} \in \mathcal{U}.$$

Sets in $\mathcal{U}$ should be thought of as big sets, or more strictly $\mathcal{U}$-*big*, with those not in $\mathcal{U}$ designated $\mathcal{U}-$*small*. The ultrafilter property means that every

[1] This takes its values in ${}^*\mathbb{R}$, and is defined just as in $\mathbb{R}$, so that $|x| = x$ if $x \geq 0$ and $|x| = -x$ if $x < 0$.

[2] A *nonprincipal ultrafilter* $\mathcal{U}$ on $\mathbb{N}$ is a collection of subsets of $\mathbb{N}$ that is closed under intersections and supersets, contains no finite sets, and for every set $A \subseteq \mathbb{N}$ has either $A \in \mathcal{U}$ or $\mathbb{N} \setminus A \in \mathcal{U}$.

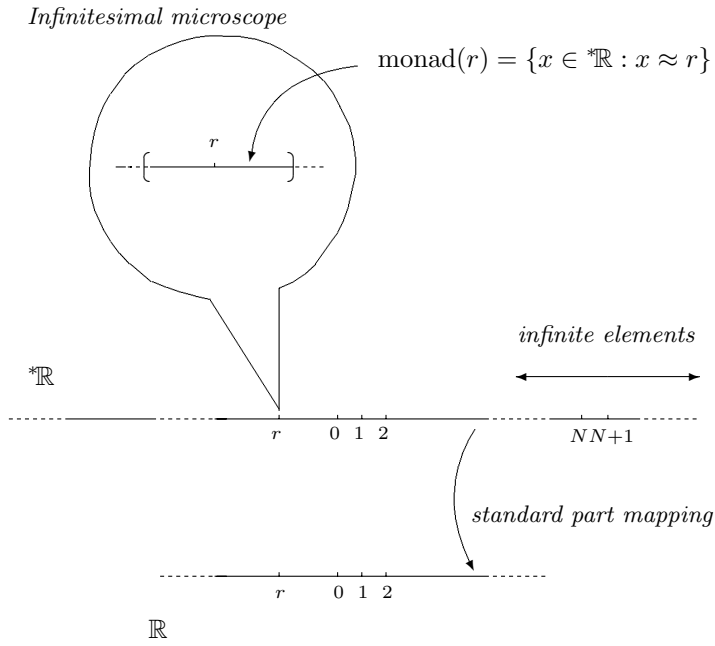
set is either $\mathcal{U}$-big (those in $\mathcal{U}$) or $\mathcal{U}$-small (those not in $\mathcal{U}$). It is convenient to use the terminology *$\mathcal{U}$-almost all* to mean "for a set A of natural numbers with $A \in \mathcal{U}$".

Using this terminology we can say that the equivalence relation $\equiv_{\mathcal{U}}$ identifies sequences (a_n) and (b_n) that agree on a $\mathcal{U}$-big set of indices n, or that agree $\mathcal{U}$-almost always.

We denote the equivalence class of a sequence (a_n) by $(a_n)_{\mathcal{U}}$ (sometimes the notation $[(a_n)]$ is used instead). The reals $\mathbb{R}$ are identified with the equivalence classes of constant sequences, so that ${}^*\mathbb{R}$ is then an extension of $\mathbb{R}$. The algebraic operations $+, \times$ and the order relation $<$ are extended to ${}^*\mathbb{R}$ pointwise (after checking that this is safe); strictly the extensions should be denoted ${}^*+, {}^*\times, {}^*<$, but there is usually no ambiguity if the * is dropped.

It is almost immediate that an example of a non-zero infinitesimal is given by $(1, \frac{1}{2}, \frac{1}{3}, \ldots)/\mathcal{U}$.

The way to picture ${}^*\mathbb{R}$ is as follows (note that some features in this diagram are yet to be explained).



The Hyperreals

With the above construction of ${}^*\mathbb{R}$ it is easy to prove:

Theorem 1.2 $({}^*\mathbb{R}, +, \times, <)$ *is an ordered field.*

Exercise Prove this! Most of the field axioms follow easily from the fact that they hold at each co-ordinate of the representing sequences. The axiom

of inverses is not quite so obvious. If $x = (a_n)_{\mathcal{U}} \neq 0$ then we could have $a_n = 0$ for some indices n. As a hint, note that nevertheless we have $a_n \neq 0$ for $\mathcal{U}$-almost all n, so we can define $y = (b_n)_{\mathcal{U}}$ with $b_n = a_n^{-1}$ for those n. For the remaining $\mathcal{U}$-small set of n define $b_n = 0$. Now show that $xy = 1$.

The axioms for an *ordered* field are proved in somewhat similar fashion.

To see what else can be said about ${}^*\mathbb{R}$, first note that *all* functions f and relations R on $\mathbb{R}$ (including unary relations – that is, subsets of $\mathbb{R}$) can be extended to ${}^*\mathbb{R}$ pointwise – with the extensions denoted by ${}^*f$ and ${}^*R$ say.[3]

As an exercise the reader might like to show that the extension ${}^*|\cdot|$ of the modulus function defined in this way is the same as that used in Definition 1.1; that is, if $x = (a_n)_{\mathcal{U}}$ and $y = {}^*|x| = (|a_n|)_{\mathcal{U}}$ then $y = x$ if $x^* \geq 0$ and $y = -x$ otherwise. Further, show that x is finite (according to Definition 1.1) if there is some $r \in \mathbb{R}$ with $|a_n| < r$ for $\mathcal{U}$-almost all n, and x is infinitesimal if, for every real $\varepsilon > 0$ we have $|a_n| < \varepsilon$ for $\mathcal{U}$-almost all n.

Important examples of extensions of relations include ${}^*\mathbb{N}, {}^*\mathbb{Z}$ and ${}^*\mathbb{Q}$, the sets of *hypernatural numbers, hyperintegers* and *hyperrationals* respectively. A hyperrational number is thus an element $x = (a_n)_{\mathcal{U}}$ with $a_n \in \mathbb{Q}$ for $\mathcal{U}$-almost all n.

It is not hard to see that the properties of functions f and relations R are transferred to (or inherited by) ${}^*f$ and ${}^*R$ – for example, if f is an injection, so is ${}^*f$, and if R is an equivalence relation then so is ${}^*R$. If $f : A \to B$ then ${}^*f : {}^*A \to {}^*B$. Moreover, connections between functions and relations are also transferred – for example ${}^*\sin^2 x + {}^*\cos^2 x = 1$ for all $x \in {}^*\mathbb{R}$. The full extent of this idea is described neatly by the *Transfer Principle* discussed below. First let us write

$$\underline{\mathbb{R}} = (\mathbb{R}, (f)_{f\in\mathcal{F}}, (R)_{R\in\mathcal{R}})$$

for the full structure with domain $\mathbb{R}$ together with every possible function and relation on it, and then write

$${}^*\underline{\mathbb{R}} = ({}^*\mathbb{R}, ({}^*f)_{f\in\mathcal{F}}, ({}^*R)_{R\in\mathcal{R}}).$$

The following fundamental result gives the complete picture as to which properties of $\underline{\mathbb{R}}$ are inherited by (or transferred to) ${}^*\underline{\mathbb{R}}$.

Theorem 1.3 (Transfer Principle) *Let φ be any first order statement. Then*

$$\varphi \quad \text{holds in} \quad \underline{\mathbb{R}} \qquad \Longleftrightarrow \qquad {}^*\varphi \quad \text{holds in} \quad {}^*\underline{\mathbb{R}}$$

A *first order* statement φ (respectively ${}^*\varphi$) is one that refers to elements of $\mathbb{R}$ (respectively ${}^*\mathbb{R}$), both fixed and variable, and to fixed relations and functions f, R (respectively ${}^*f, {}^*R$). First order statements can use the usual

[3] By the pointwise extension of a binary relation $R \subset \mathbb{R} \times \mathbb{R}$, say, we mean that $((a_n)_{\mathcal{U}}, (b_n)_{\mathcal{U}}) \in {}^*R \Leftrightarrow (a_n, b_n) \in R$ for $\mathcal{U}$-almost all n; so ${}^*R \subset {}^*\mathbb{R} \times {}^*\mathbb{R}$. It is easy to see that this is equivalent to defining ${}^*R$ using a pointwise extension of the characteristic function – i.e. $\chi_{{}^*R}((a_n)_{\mathcal{U}}, (b_n)_{\mathcal{U}}) = (\chi_R(a_n, b_n))_{\mathcal{U}}$.

logical connectives of mathematics, namely *and* (symbolically $\wedge$), *or* ($\vee$), *implies* ($\rightarrow$) and *not* ($\neg$). Moreover, we can quantify over *elements* ($\forall x$, $\exists y$) but not over relations or functions (so $\forall f, \exists R$ are not allowed). Here are some illustrations of this.

1. Density of the rationals in the reals.

The density of the rationals in the reals can be expressed by a first order statement φ that is a formal version of the following.

Between every two distinct reals there is a rational.

We could for example take φ as the statement

$$\forall x \forall y \, (x < y \rightarrow \exists z (z \in \mathbb{Q} \wedge (x < z < y)))$$

The transfer principle tells us that ${}^*\varphi$ holds in ${}^*\underline{\mathbb{R}}$ which means that the hyperrationals are dense in the hyperreals.

2. Discreteness of the ordering of the integers.

This can be expressed by a first order statement ψ which is a formal version of the following.

Every $n \in \mathbb{Z}$ has an immediate predecessor and successor in $\mathbb{Z}$.

The Transfer Principle tells us that ${}^*\psi$ holds in ${}^*\underline{\mathbb{R}}$ which means that

Every $n \in {}^*\mathbb{Z}$ has an immediate predecessor and successor in ${}^*\mathbb{Z}$.

Thus the discreteness of $\mathbb{Z}$ is inherited by ${}^*\mathbb{Z}$.

The reader is invited to check that the immediate predecessor and successor of a hyperinteger $(m_n)_{\mathcal{U}} \in {}^*\mathbb{Z}$ are given by $(m_n \mp 1)_{\mathcal{U}}$. Likewise the density of the hyperrationals can be established quite easily from first principles (if $x = (a_n)_{\mathcal{U}}$ and $y = (b_n)_{\mathcal{U}}$ then take $z = (c_n)_{\mathcal{U}}$ with $a_n < c_n < b_n$ where possible). The proof of the Transfer Principle is just a generalisation of the procedure involved in a direct verification of these two examples. The Transfer Principle itself avoids the need to verify properties of ${}^*\underline{\mathbb{R}}$ on an *ad hoc* basis, and instead gives us all properties from the beginning.

A key result that allows us to get back to $\mathbb{R}$ from ${}^*\mathbb{R}$ (and extends to more general topological situations) is the following (recall the definition 1.1 of a *finite* hyperreal).

Theorem 1.4 (Standard Part Theorem) *If $x \in {}^*\mathbb{R}$ is finite, then there is a unique $r \in \mathbb{R}$ such that $x \approx r$; i.e. any finite hyperreal x is uniquely expressible as $x = r + \delta$ with r a* standard *real and δ infinitesimal.*

Proof Put $r = \sup\{a \in \mathbb{R} : a \leq x\} = \sup A$, say. The set A is nonempty and bounded above (in $\mathbb{R}$) since x is finite, and so the least upper bound r exists. It is routine to check that $|x - r| < \varepsilon$ for every real $\varepsilon > 0$. $\square$

Definition 1.5 (Standard Part) If x is a finite hyperreal the unique real $r \approx x$ is called the *standard part* of x.

For a finite hyperreal $x \in {}^*\mathbb{R}$ there are two notations (both useful) for the standard part of x:

$$^\circ x = \mathrm{st}(x) = \text{the } \textit{standard part} \text{ of } x.$$

On occasions, when considering extended real valued functions (with values in $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$), it is convenient to write $^\circ x = \pm\infty$ if x is positive (resp. negative) infinite.

Remark The Standard Part Theorem is equivalent to the completeness of $\mathbb{R}$.

The next two theorems illustrate the way in which real analysis develops using the additional structure of ${}^*\mathbb{R}$. For the sake of completeness we give brief proofs that provide a flavour of the nonstandard methodology, and especially the use of the Transfer Principle.

For a full account of the development of real analysis using infinitesimals, see any of the references [30, 54, 47, 56, 58, 60, 69].

For the following, note that since a sequence $s = (s_n)_{n\in\mathbb{N}}$ of reals is just a function $s : \mathbb{N} \to \mathbb{R}$, its nonstandard extension ${}^*s = ({}^*s_n)_{n\in{}^*\mathbb{N}}$ is simply a function ${}^*s : {}^*\mathbb{N} \to {}^*\mathbb{R}$.

Theorem 1.6 *Let (s_n) be a sequence of real numbers and let $l \in \mathbb{R}$. Then*

$$s_n \to l \text{ as } n \to \infty \quad \Longleftrightarrow \quad {}^*s_K \approx l \text{ for all infinite } K \in {}^*\mathbb{N}.$$

Proof Suppose first that $s_n \to l$, and fix infinite $K \in {}^*\mathbb{N}$. We have to show that $|{}^*s_K - l| < \varepsilon$ for all real $\varepsilon > 0$.

For any such ε there is a number $n_0 \in \mathbb{N}$ such that the following holds in $\mathbb{R}$:

$$\forall n \in \mathbb{N}[n \geq n_0 \to |s_n - l| < \varepsilon]$$

The Transfer Principle now tells us that

$$\forall N \in {}^*\mathbb{N}[N \geq n_0 \to |{}^*s_N - l| < \varepsilon]$$

is true in ${}^*\mathbb{R}$. In particular taking $N = K$ we see that $|{}^*s_K - l| < \varepsilon$ as required.

Conversely, suppose that ${}^*s_K \approx l$ for all infinite $K \in {}^*\mathbb{N}$. Then, for any given real $\varepsilon > 0$, we have

$$\exists K \in {}^*\mathbb{N}\ \forall N \in {}^*\mathbb{N}[N \geq K \to |{}^*s_N - l| < \varepsilon]$$

The Transfer Principle applied to this statement shows that in $\mathbb{R}$:

$$\exists k \in \mathbb{N} \forall n \in \mathbb{N}[n \geq k \to |s_n - l| < \varepsilon]$$

Taking n_0 to be any such k proves that $s_n \to l$. □

For the next result note that if f is a real function defined on the open interval $]a, b[$ then ${}^*f$ is defined on the hyperreal interval $]a, b[= \{x \in {}^*\mathbb{R} : a < x < b\}$, and takes values in ${}^*\mathbb{R}$.

Theorem 1.7 *Let $c \in]a, b[$ (where $a, b, c \in \mathbb{R}$) and $f :]a, b[\to \mathbb{R}$. Then*

$$f \text{ is continuous at } c \iff {}^*f(z) \approx f(c) \text{ whenever } z \approx c \text{ in } {}^*\mathbb{R}.$$

Proof The proof is very similar to that of Theorem 1.6. Suppose first that f is continuous at c, and fix a hyperreal $z \approx c$. We have to show that $|{}^*f(z) - f(c)| < \varepsilon$ for all real $\varepsilon > 0$.

For any such ε there is a number $0 < \delta \in \mathbb{R}$ such that the following holds in $\mathbb{R}$:

$$\forall x[|x - c| < \delta \to |f(x) - f(c)| < \varepsilon]$$

The Transfer Principle now tells us that

$$\forall \mathrm{X}[|\mathrm{X} - c| < \delta \to |{}^*f(\mathrm{X}) - f(c)| < \varepsilon]$$

is true in ${}^*\mathbb{R}$. In particular taking $X = z$ we see that $|{}^*f(z) - f(c)| < \varepsilon$ as required.

Conversely, suppose that $|{}^*f(z) - f(c)| \approx 0$ for all $z \approx c$ in ${}^*\mathbb{R}$. Let a real $\varepsilon > 0$ be given. Then taking Y to be any positive infinitesimal the following holds in ${}^*\mathbb{R}$:

$$\exists Y \forall \mathrm{X}[|\mathrm{X} - c| < Y \to |{}^*f(\mathrm{X}) - f(c)| < \varepsilon]$$

The Transfer Principle applied to this statement gives, in $\mathbb{R}$:

$$\exists y \forall x[|x - c| < y \to |f(x) - f(c)| < \varepsilon]$$

Taking δ to be any such y shows that f is continuous at c as required. □

Before moving to the next section, it should be pointed out that there are several other ways to construct the hyperreals. Moreover, the conventional terminology is misleading in that different constructions do not necessarily give isomorphic structures. All versions of the hyperreals however obey the Transfer Principle, and this is all that is needed to do basic nonstandard real analysis. Indeed, one perfectly workable approach to the subject is an axiomatic one, which merely specifies that ${}^*\mathbb{R}$ is an extension of $\mathbb{R}$ that obeys the Transfer Principle. (This approach would be parallel to a development of real analysis that proceeds without being concerned with any particular construction of $\mathbb{R}$, using only the assumption that $\mathbb{R}$ is a complete ordered field.)

1.2.2 The nonstandard universe

To use Robinson's ideas beyond the realm of real analysis, it is necessary to repeat the construction of ${}^*\mathbb{R}$ for any mathematical object $\mathcal{M}$ that might be needed, giving a nonstandard version ${}^*\mathcal{M}$ of $\mathcal{M}$ that contains ideal elements (such as infinitesimals in the case of ${}^*\mathbb{R}$). $\mathcal{M}$ could be a group, ring, measure space, metric space or *any* mathematical object, however complicated.

Rather than construct each nonstandard extension ${}^*\mathcal{M}$ as required, it is more economical to construct at the outset a nonstandard version ${}^*\mathbb{V}$ of a working portion of the mathematical universe $\mathbb{V}$ that contains each object $\mathcal{M}$ that might be needed. Then ${}^*\mathbb{V}$ will contain ${}^*\mathcal{M}$ for every $\mathcal{M} \in \mathbb{V}$. Such a construction has the additional advantage that the corresponding Transfer Principle preserves connections *between* structures as well as their intrinsic properties.

Here, briefly, is the way it works. First, for most mathematical practice, an adequate portion of the mathematical universe is the *superstructure over* $\mathbb{R}$, denoted by $\mathbb{V} = \mathbb{V}(\mathbb{R})$, defined as follows.

$$\begin{aligned} V_0(\mathbb{R}) &= \mathbb{R} \\ V_{n+1}(\mathbb{R}) &= V_n(\mathbb{R}) \cup \mathcal{P}(V_n(\mathbb{R})), \quad n \in \mathbb{N} \end{aligned}$$

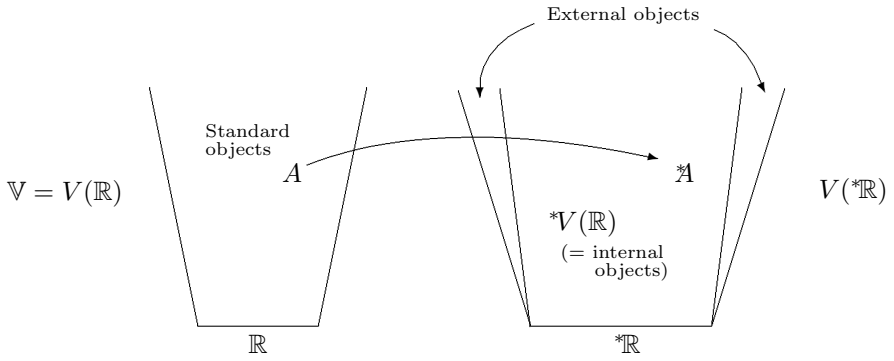
and

$$\mathbb{V} = V(\mathbb{R}) = \bigcup_{n\in\mathbb{N}} V_n(\mathbb{R}).$$

(If $V(\mathbb{R})$ is not big enough to contain all the objects[4] required, simply replace the starting set $\mathbb{R}$ by a suitable larger set S, giving $\mathbb{V} = V(S)$.)

The next step is to construct a mapping ${}^* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ which associates to each object $\mathcal{M} \in \mathbb{V}$ a nonstandard extension ${}^*\mathcal{M} \in V({}^*\mathbb{R})$. Roughly, we have $\mathcal{M} \subset {}^*\mathcal{M}$ with ${}^*\mathcal{M} \setminus \mathcal{M}$ consisting of "ideal" or "nonstandard" elements. For example ${}^*\mathbb{N} \setminus \mathbb{N}$ consists of infinite (hyper)natural numbers; if $\mathcal{M}$ is an infinite dimensional Hilbert space $\mathbf{H}$ together with its finite dimensional subspaces then ${}^*\mathcal{M}$ will contain some infinite hyperfinite dimensional subspaces.

The way to visualise the resulting *nonstandard universe* is as follows.



The Nonstandard Universe

The nonstandard universe is in fact the collection

[4] We are now taking the view that all mathematical objects are sets.

$$^*\mathbb{V} = \{x : x \in {}^*\mathcal{M} \text{ for some } \mathcal{M} \in \mathbb{V}\}$$

consisting of all new and old members of sets in $\mathbb{V}$. Although $^*\mathbb{V} \subset V(^*\mathbb{R})$, it is crucial to realise that $^*\mathbb{V}$ is **not** the same as $V(^*\mathbb{R})$. Sets in $^*\mathbb{V}$ are known as *internal sets.*

One way[5] to construct $^*\mathbb{V}$ is by means of an ultrapower

$$\mathbb{V}^{\mathbb{N}}/\mathcal{U}$$

although there is a little more work to do (compared to the corresponding construction of $^*\mathbb{R}$). The set membership relation $\in$ that gives the structure $(\mathbb{V}, \in)$, when extended pointwise to the ultrapower $\mathbb{V}^{\mathbb{N}}/\mathcal{U}$, gives a "pseudo-membership" relation E, say, resulting in the structure

$$(\mathbb{V}^{\mathbb{N}}/\mathcal{U}, E).$$

It is then necessary to take the "Mostowski collapse" of this structure, which constructs simultaneously the collection $^*\mathbb{V}$ and an injection

$$i : (^*\mathbb{V}, \in) \to (\mathbb{V}^{\mathbb{N}}/\mathcal{U}, E).$$

Although i is not surjective, its range includes the equivalence class of each constant sequence, and then $^*\mathcal{M}$ is defined by

$$^*\mathcal{M} = i^{-1}((\mathcal{M}, \mathcal{M}, \mathcal{M}, \ldots \mathcal{M})/\mathcal{U}).$$

The key property of the nonstandard universe $^*\mathbb{V}$ is a Transfer Principle which again indicates precisely which properties of the superstructure $\mathbb{V}$ are inherited by $^*\mathbb{V}$.

Theorem 1.8 (The Transfer Principle) *Suppose that φ is a bounded quantifier statement. Then φ holds in $\mathbb{V}$ if and only if $^*\varphi$ holds in $^*\mathbb{V}$.*

A *bounded quantifier statement* (*bqs*) is simply a statement of mathematics that can be written in such a way that all quantifiers range over a prescribed set. That is, we have subclauses such as $\forall x \in A$ and $\exists y \in B$ but not unbounded quantifiers such as $\forall x$ and $\exists y$. Most quantifiers in mathematical practice are bounded (often only implicitly in exposition). A *bqs* φ may also contain fixed sets $\mathcal{M}$ from $\mathbb{V}$, which will be replaced in $^*\varphi$ by $^*\mathcal{M}$.

Members of internal sets are internal (this follows easily from the construction) and since the sets $^*\mathcal{M}$ are also internal, it follows that the information we obtain from the Transfer Principle is entirely about *internal sets.* To illustrate, the Transfer Principle tells us that any *internal* bounded subset of $^*\mathbb{R}$

[5] This sketch of a construction of $^*\mathbb{V}$ can be skipped without any loss – it is included to show that a nonstandard universe is a very down-to-earth and non-mysterious mathematical construct.

has a least upper bound, whereas this can fail for external[6] sets. For example, the set $\mathbb{N}$ is a subset of ${}^*\mathbb{R}$ that is bounded (by any infinite hyperreal) but has no least upper bound – from which we deduce that $\mathbb{N}$ is *external.* Incidentally this demonstrates that there actually are external sets – i.e. $\mathbb{V}({}^*\mathbb{R}) \setminus {}^*\mathbb{V} \neq \boldsymbol{\Omega}$.

An easy application of the Transfer Principle gives the following very useful properties.

Proposition 1.9 *Let $A \subseteq {}^*\mathbb{R}$ be an internal set.*

(a) *(Overflow) If A contains arbitrarily large finite numbers then it also contains an infinite number;*

(b) *(Underflow) If A contains arbitrarily small positive infinite[7] numbers then it contains a positive finite number.*

Taking reciprocals gives a corresponding pair of principles for the set of infinitesimals.

As with ${}^*\mathbb{R}$ it is possible (and quite convenient) to take an axiomatic approach to ${}^*\mathbb{V}$, which simply postulates the existence of a set ${}^*\mathbb{V}$ and a mapping ${}^* : \mathbb{V} \to {}^*\mathbb{V}$ that obeys the Transfer Principle. For most purposes (and certainly the construction of Loeb measures) one further assumption is needed, which we now discuss.

1.2.3 $\aleph_1$-saturation

A nonstandard universe constructed as a countable ultrapower has an additional property called $\aleph_1$-saturation, which we highlight here because of its importance.

Definition 1.10 A nonstandard universe ${}^*\mathbb{V}$ is said to be $\aleph_1$*-saturated* if the following holds*:*

if $(A_m)_{m\in\mathbb{N}}$ is a countable decreasing sequence of *internal* sets with each $A_m \neq \boldsymbol{\Omega}$, then $\bigcap_{m\in\mathbb{N}} A_m \neq \boldsymbol{\Omega}$.

Theorem 1.11 *A nonstandard universe ${}^*\mathbb{V}$ constructed as a countable ultrapower is $\aleph_1$-saturated.*

Proof (Sketch) Each set A_m is represented by a sequence of standard sets $(X_{m,n})_{n\in\mathbb{N}}$. Since each A_m is nonempty and the sequence is decreasing, then for $\mathcal{U}$-almost all[8] n we have $X_{m+1,n} \subseteq X_{m,n}$ and $X_{m,n} \neq \boldsymbol{\Omega}$. By a systematic modification of the sets $X_{m,n}$ on a $\mathcal{U}$-small[9] set of indices n we may assume that $X_{m+1,n} \subseteq X_{m,n}$ and $X_{m,n} \neq \boldsymbol{\Omega}$ for all n and m. Now pick $x_n \in X_{n,n}$

[6] an *external set* is one that is not internal

[7] that is, for every positive infinite $x \in {}^*\mathbb{R}$ there is an element $a \in A$ with a infinite and $a < x$

[8] that is, the set $\{n : X_{m+1,n} \subseteq X_{m,n}\}$ belongs to $\mathcal{U}$.

[9] i.e. a set that is not in the ultrafilter $\mathcal{U}$.

and let $y = (x_n)_{\mathcal{U}}$ be the element represented by this sequence. Then $y \in A_m$ for every m since $x_n \in X_{m,n}$ for $n \geq m$. $\square$

$\aleph_1$-saturation is a kind of compactness property that is essential for the Loeb measure construction. **For the rest of these lectures we assume that ${}^*\mathbb{V}$ is a nonstandard universe that is $\aleph_1$-saturated.**

It is possible to build nonstandard universes with stronger saturation type properties, by an extension of the techniques discussed above. These are needed in some applications of Loeb measures involving topological spaces that do not have a countable sub-base, and in other "non-separable" mathematical applications.

An equivalent and very useful formulation of $\aleph_1$-saturation, known as *countable comprehension*, goes as follows.

Countable comprehension Given any sequence $(A_n)_{n\in\mathbb{N}}$ of internal subsets of an internal set A, there is an *internal* sequence[10] $(A_n)_{n\in{}^*\mathbb{N}}$ of subsets of A that extends the original sequence.

To see that $\aleph_1$-saturation implies countable comprehension, apply $\aleph_1$-saturation to the sets B_m consisting of internal sequences $(C_n)_{n\in{}^*\mathbb{N}}$ with $C_n = A_n$ for $n \leq m$. The reverse implication is proved using the overflow principle. (The reader may like to try proving this as an exercise.)

1.2.4 Nonstandard topology

We gather together here some of the basic nonstandard topological notions that will be referred to later.

First, we note that the idea of being infinitely close generalises to any topological space, extending the idea of a *monad.* Recall that for $a \in \mathbb{R}$ the *monad* of a is the set $\text{monad}(a) = \{x \in {}^*\mathbb{R} : x \approx a\}$.

More generally we have:

Definition 1.12 Let $(X, \mathcal{T})$ be a topological space.

(i) For $a \in X$ the *monad* of a is

$$\text{monad}(a) = \bigcap_{a\in U\in\mathcal{T}} {}^*U.$$

(ii) If $x \in {}^*X$, we write $x \approx a$ to mean $x \in \text{monad}(a)$. (Note that in general this is not a symmetric relationship.)

(iii) $x \in {}^*X$ is *nearstandard* if $x \approx a$ for some $a \in X$.

(iv) $\text{ns}(Y)$ is the set of nearstandard points in Y, for any $Y \subseteq {}^*X$.

(v) $\text{st}(Y) = \{a \in X : x \approx a \text{ for some } x \in Y\}$; this is called the *standard part* of Y.

[10] that is, an internal function with domain ${}^*\mathbb{N}$.

The idea of the pointwise standard part mapping for ${}^*\mathbb{R}$ generalizes to Hausdorff spaces because of the next result.

Proposition 1.13 *A topological space X is Hausdorff if and only if*

$$\text{monad}(a) \cap \text{monad}(b) = \boldsymbol{\Omega} \quad \text{for} \quad a \neq b, \quad a, b \in X.$$

Proof An easy exercise. □

This means that for Hausdorff spaces we can define the *standard part* mapping

$$\text{st} : \text{ns}({}^*X) \to X$$

by

$$\text{st}(x) = \text{the unique } a \in X \text{ with } a \approx x.$$

The following notation is often used:

$$^\circ x = \text{st}(x).$$

If necessary we write st_X or $\text{st}_{\mathcal{T}}$ to denote the space or topology concerned.

The following is another important notion that plays a key rôle in constructing solutions to differential equations of all kinds.

Definition 1.14 Suppose that Y is a subset of ${}^*X$ for some topological space X, and $F : {}^*X \to {}^*\mathbb{R}$ is internal. Then F is said to be *S-continuous on* Y if for all $x, y \in Y$ we have

$$x \approx y \Longrightarrow F(x) \approx F(y).$$

The importance of this notion is seen in the following result.

Theorem 1.15 *If $F : {}^*\mathbb{R} \to {}^*\mathbb{R}$ is S-continuous on an interval ${}^*[a, b]$ for real a, b, and $F(x)$ is finite for some $x \in {}^*[a, b]$, then the standard function defined on $[a, b]$ by*

$$f(t) = {}^\circ F(t)$$

is continuous, and ${}^*f(\tau) \approx F(\tau)$ for all $\tau \in {}^*[a, b]$.

Remark This theorem shows that S-continuous functions in ${}^*C[a, b]$are precisely those that are nearstandard in the uniform topology on $C[a, b]$, and the function f defined above is the standard part ${}^\circ F$ for this topology.

One final result from general nonstandard topology that we will need is:

Proposition 1.16 *Let $(X, \mathcal{T})$ be separable, Hausdorff. Suppose that $Y \subseteq {}^*X$ is internal, and $A \subseteq X$. Then*

(a) $\text{st}(Y)$ *is closed,*
(b) *if X is regular and $Y \subseteq \text{ns}({}^*X)$ then $\text{st}(Y)$ is compact,*
(c) $\text{st}({}^*A) = \overline{A}$ *(the closure of A),*
(d) *if X is regular, then A is relatively compact iff ${}^*A \subseteq \text{ns}({}^*X)$.*

Remark The condition that X should be separable in Proposition 1.16 can be omitted if the nonstandard model has more saturation – namely κ-saturation (see the Remark at the end of the previous section), where the topology on X has a base of cardinality κ. However, in all our applications the relevant spaces X are separable, and so $\aleph_1$-saturation (which we have in our model) is sufficient.

1.3 Construction of Loeb Measures

A *Loeb measure* is a measure constructed from a nonstandard measure by the following construction of Peter Loeb [71]. We confine our attention in these lectures mainly to finite (or bounded) Loeb measures.

Suppose that an internal set Ω and an internal algebra $\mathcal{A}$ of subsets of Ω are given, and suppose further that μ is a finite internal finitely additive measure on $\mathcal{A}$. This means that μ is an internal mapping

$$\mu : \mathcal{A} \to {}^*[0, \infty)$$

with $\mu(A \cup B) = \mu(A) + \mu(B)$ for disjoint A, $B \in \mathcal{A}$, and that $\mu(\Omega)$ is finite.[11] Thus $\mu(A)$ is finite for each $A \in \mathcal{A}$, so we may define the mapping

$$\circ\mu : \mathcal{A} \to [0, \infty)$$

by $\circ\mu(A) = \circ(\mu(A))$. Clearly $\circ\mu$ is finitely additive, so that $(\Omega, \mathcal{A}, \circ\mu)$ is a *standard finitely additive measure space*.

In general this is *not* a measure space, because $\mathcal{A}$ is not σ-additive unless $\mathcal{A}$ is finite.

Nevertheless, if $(A_n)_{n\in\mathbb{N}}$ is a family of sets from $\mathcal{A}$, then the set $\bigcup_{n\in\mathbb{N}} A_n$ is almost in $\mathcal{A}$. It differs from a set in $\mathcal{A}$ by a *null* set (a notion to be defined shortly); see the Key Lemma (Lemma 1.19) and its corollary below. This is what lies at the heart of the following fundamental result proved by Loeb.

Theorem 1.17 *There is a unique σ-additive extension of $\circ\mu$ to the σ-algebra $\sigma(\mathcal{A})$ generated by $\mathcal{A}$. The completion of this measure is the* Loeb measure corresponding to μ, *denoted μ_L and the completion of $\sigma(\mathcal{A})$ is the* Loeb σ-algebra, *denoted by $L(\mathcal{A})$.*

Proof For a quick proof we can apply Caratheodory's extension theorem. It is only necessary to check σ-additivity of $\circ\mu$ on $\mathcal{A}$. Suppose that $(A_n)_{n\in\mathbb{N}}$ is a sequence of pairwise disjoint sets from $\mathcal{A}$ such that

$$A = \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}.$$

[11] It also means of course that all sets in $\mathcal{A}$ are internal

By $\aleph_1$-saturation (applied to the decreasing sequence of sets $A \setminus \bigcup_{n=1}^{m} A_n$) there is $m \in \mathbb{N}$ such that

$$\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n=1}^{m} A_n.$$

So $A_k = \boldsymbol{\Omega}$ for $k > m$, and

$$^\circ\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = {}^\circ\mu\left(\bigcup_{n=1}^{m} A_n\right) = \sum_{n=1}^{m} {}^\circ\mu(A_n) = \sum_{n\in\mathbb{N}} {}^\circ\mu(A_n),$$

using finite additivity. Caratheodory's theorem (see [88] for example) now gives the result. □

It is quite straightforward and rather more illuminating to prove Loeb's theorem from "first principles" and here is one way to proceed – based around the idea of a *Loeb null set*. (See [29] for full details of this approach.)

Definition 1.18 Let $B \subseteq \Omega$ (not necessarily internal). We say that B is a *Loeb null set* if for each real $\varepsilon > 0$ there is a set $A \in \mathcal{A}$ with $B \subseteq A$ and $\mu(A) < \varepsilon$.

The following result makes it clear that $\mathcal{A}$ is almost a σ-algebra.

Lemma 1.19 (Key Lemma) *Let $(A_n)_{n\in\mathbb{N}}$ be an increasing family of sets, with each A_n in $\mathcal{A}$, and let $B = \bigcup_{n\in\mathbb{N}} A_n$. Then there is a set $A \in \mathcal{A}$ such that*

(a) $B \subseteq A$;
(b) $^\circ\mu(A) = \lim_{n\to\infty} {}^\circ\mu(A_n)$;
(c) $A \setminus B$ *is null.*

Proof Let $\alpha = \lim_{n\to\infty} {}^\circ\mu(A_n)$. For each finite n,

$$\mu(A_n) \;\le\; {}^\circ\mu(A_n) + \frac{1}{n} \;\le\; \alpha + \frac{1}{n}.$$

Now, using $\aleph_1$-saturation, take an increasing *internal* sequence $(A_n)_{n\in{}^*\mathbb{N}}$ of sets in $\mathcal{A}$ extending the sequence $(A_n)_{n\in\mathbb{N}}$. Overflow gives an infinite N such that

$$\mu(A_N) \le \alpha + \frac{1}{N}.$$

Let $A = A_N$. Then (a) holds because $A \supseteq A_n$ for each finite n. Moreover, $\mu(A_n) \le \mu(A)$ for each finite n, so $^\circ\mu(A_n) \le {}^\circ\mu(A) \le \alpha$, giving $^\circ\mu(A) = \alpha$, which is (b). Moreover, $^\circ\mu(A \setminus A_n) = {}^\circ\mu(A) - {}^\circ\mu(A_n) \to 0$. Now $A \setminus B \subseteq A \setminus A_n$ so $A \setminus B$ is null. □

From this Key Lemma, it is clear that $\mathcal{A}$ is almost a σ-algebra – and in fact it is a σ-algebra modulo null sets. The following makes this precise.

Definition 1.20 (i) Let $B \subseteq \Omega$. We say that B is *Loeb measurable* if there is a set $A \in \mathcal{A}$ such that $A \Delta B$[12] is Loeb null. Denote the collection of all Loeb measurable sets by $L(\mathcal{A})$.

(ii) For $B \in L(\mathcal{A})$ define

$$\mu_L(B) = {}^\circ\mu(A)$$

for any $A \in \mathcal{A}$ with $A \Delta B$ null, and call $\mu_L(B)$ the *Loeb measure* of B.

It is then quite straightforward to prove:

Theorem 1.21 *$L(\mathcal{A})$ is a σ-algebra, and μ_L is a complete (σ-additive) measure on $L(\mathcal{A})$.*

The measure space $\boldsymbol{\Omega} = (\Omega, L(\mathcal{A}), \mu_L)$ is called the *Loeb space* given by $(\Omega, \mathcal{A}, \mu)$, and $L(\mathcal{A})$ is called the *Loeb algebra*. Of course $L(\mathcal{A})$ depends on both $\mathcal{A}$ and μ, so strictly we should write $L(\mathcal{A}, \mu)$, but usually it is clear which measure is intended. If $\mu(\Omega) = 1$ then $\boldsymbol{\Omega}$ is a *Loeb probability space* and μ_L is the *Loeb probability measure* given by μ.

The following are alternative characterisations of Loeb measurable sets, and are often taken as the fundamental definition (see [3], [20] or [69] for example). First some definitions are required.

Definition 1.22 Let $B \subseteq \Omega$ (not necessarily internal).

(i) B is *μ-approximable* if for every real $\varepsilon > 0$ there are sets $A, C \in \mathcal{A}$ with $A \subseteq B \subseteq C$ and $\mu(C \setminus A) < \varepsilon$.

(ii) The *inner* and *outer* Loeb measure of B, $\underline{\mu}(B)$ and $\overline{\mu}(B)$ are given by

$$\underline{\mu}(B) = \sup\{{}^\circ\mu(A) : \ A \subseteq B, \ A \in \mathcal{A}\}$$

$$\overline{\mu}(B) = \inf\{{}^\circ\mu(A) : \ A \supseteq B, \ A \in \mathcal{A}\}$$

Then we have

Theorem 1.23 *The following are equivalent:*

(a) *B is Loeb measurable.*

(b) *B is μ-approximable.*

(c) $\overline{\mu}(B) = \underline{\mu}(B)$.

[12] $A \Delta B$ is the symmetric difference $(A \setminus B) \cup (B \setminus A)$

Loeb counting measure

For a simple illustration of the Loeb construction (but one which has far reaching applications) consider the *Loeb counting measure*, as follows. Let $\Omega = \{1, 2, \ldots, N\}$ where $N \in {}^*\mathbb{N} \setminus \mathbb{N}$, so that Ω is a infinite *hyperfinite* set (necessarily internal), and let ν be the counting probability measure on Ω, defined by

$$\nu(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{N}$$

for $A \in {}^*\mathcal{P}(\Omega) = \mathcal{A}$, say.[13] Here $|A|$ denotes the number[14] of elements in A. Note that ${}^*\mathcal{P}(\Omega)$ is a proper subset of $\mathcal{P}(\Omega)$, since, for example, the set $\mathbb{N} \in \mathcal{P}(\Omega) \setminus {}^*\mathcal{P}(\Omega)$, which in turn shows that $\mathcal{A}$ is *not* a σ-algebra. The *Loeb counting measure* ν_L is the completion of the extension to $\sigma(\mathcal{A})$ of the finitely additive measure ${}^\circ\nu$.

1.3.1 Example: Lebesgue measure

A first simple application of Loeb measure is an intuitive construction of Lebesgue measure. First we define the *hyperfinite (time)*[15] *line* $\mathbf{T}$ corresponding to the interval $[0, 1]$**.**

Definition 1.24 Fix $N \in {}^*\mathbb{N} \setminus \mathbb{N}$ and let $\Delta t = N^{-1}$. The *hyperfinite time line* (based on Δt, for the interval $[0, 1]$) is the set

$$\mathbf{T} = \{0, \Delta t, 2\Delta t, 3\Delta t, \ldots, 1 - \Delta t\}.$$

(In applications hyperfinite time lines may be taken with different end points, according to need.)

We will use sanserif symbols t, s for elements of $\mathbf{T}$ to distinguish them from those in $[0, 1]$.

Theorem 1.25 *Let ν_L be the Loeb counting measure on the hyperfinite time line $\mathbf{T}$. Define*

(i) $\mathcal{M} = \{B \subseteq [0, 1] : \mathrm{st}_{\mathbf{T}}^{-1}(B)$ is Loeb measurable$\}$, *where* $\mathrm{st}_{\mathbf{T}}^{-1}(B) = \{\mathsf{t} \in \mathbf{T} : {}^\circ\mathsf{t} \in B\}$.

(ii) $\lambda(B) = \nu_L(\mathrm{st}_{\mathbf{T}}^{-1}(B))$ *for* $B \in \mathcal{M}$.

Then $([0, 1], \mathcal{M}, \lambda)$ is Lebesgue measure (i.e. $\mathcal{M}$ is the Lebesgue completion of the Borel sets $\mathcal{B}[0, 1]$, and $\lambda(B)$ is the Lebesgue measure of $B \in \mathcal{M}$.)

[13] An application of the Transfer Principle tells us that this is the collection of all *internal* subsets of Ω.

[14] The Transfer Principle tells us that for *internal* subsets A of Ω there is a unique $M \in {}^*\mathbb{N}$, $M \leq N$, such that there is an internal bijection $F : \{1, 2, \ldots, M\} \to A$ – and this M is what is meant by $|A|$. Equivalently, $|\cdot|$ is the extension to ${}^*\mathbb{V}$ of the standard function $|\cdot|$ that gives the cardinality of finite sets.

[15] This has become the conventional terminology for this discrete representation of the interval [0,1] when it is used to represent time.

Proof (Sketch) It is routine to check that $\mathcal{M}$ is a σ-algebra that contains each standard interval $[a,b]$ (since $\mathrm{st}_{\mathbf{T}}^{-1}([a,b]) = \bigcap_{n\in\mathbb{N}} \left({}^*[a-\frac{1}{n}, b+\frac{1}{n}] \cap \mathbf{T}\right)$, which is a countable intersection of internal sets), and that λ is a complete probability measure on $\mathcal{M}$. Showing that λ is translation invariant and $\lambda([a,b]) = b-a$ is straightforward, so that $([0,1], \mathcal{M}, \lambda)$ is an extension of Lebesgue measure. Now take $B \in \mathcal{M}$, and an inner approximation $A \subseteq \mathrm{st}_{\mathbf{T}}^{-1}(B)$ with A internal. Then the set $\mathrm{st}(A)$ is a closed inner approximation of B, and this suffices to show that B is Lebesgue measurable. □

This result is a particular case of a general theorem of Anderson [6] that shows how any Radon measure on a Hausdorff space can be represented by a hyperfinite Loeb counting measure. A famous example of this is Anderson's representation of Wiener measure, below. A less well known but very pleasant example is David Ross' very intuitive construction of Haar measure[16], as follows (taken from [85, 87]).

1.3.2 Example: Haar measure

Let G be a compact group, and take an internal infinitesimal neighbourhood[17] V of 1. Take a minimal *open cover Ω of ${}^*G$ consisting of sets that are translates of V. So $\Omega = \{V_1, \ldots, V_N\}$ say with each $V_i = g_i V$ for some $g_i \in {}^*G$. Let ν_L be the Loeb counting probability measure on Ω. For Borel sets $B \subseteq G$ define

$$m(B) = \nu_L(\mathrm{st}_{\Omega}^{-1}(B))$$

where $\mathrm{st}_{\Omega} : \Omega \to G$ is the generalisation of the standard part mapping[18] to this context. Then m is Haar measure on G.

To see this, first it routine to show that m is a Borel probability measure on G; the other required property is that m should be translation invariant – that is, $m(B) = m(gB)$ for each $B \in \mathcal{B}$ and $g \in G$. It is sufficient to show that $m(B) \leq m(gB)$, and for this take an internal set $A \subseteq \mathrm{st}_{\Omega}^{-1}(B)$. Let $C = \{V_j : V_j \cap gV_k \neq \boldsymbol{\Omega} \text{ for some } V_k \in A\}$ and note that $C \subseteq \mathrm{st}_{\Omega}^{-1}(gB)$. It is easy to check that the collection $(\Omega \setminus A) \cup g^{-1}C$ is a cover of ${}^*G$ by sets that are translates of V, so by minimality of the collection Ω this gives $|C| \geq |A|$. Thus $m(gB) \geq m(B)$ as required. □

1.3.3 Example: Wiener measure

Perhaps the best known measure construction using Loeb measure theory is Anderson's construction [5] of Wiener measure, which we now describe. Recall

[16] *Haar measure* on a compact group is the unique probability measure that is invariant under multiplication by group elements.

[17] This means that $V \subset {}^*U$ for each open neighbourhood U of 1.

[18] Actually we have the mapping $\mathrm{st}_G : {}^*G \to G$, but since V is an infinitesimal neighbourhood, the set $\mathrm{st}_G(V_i)$ is a singleton for each i, so it makes sense to define $\mathrm{st}_{\Omega} : \Omega \to G$.

that Wiener measure W on $\mathcal{C} = C_0[0,1]$ (the set of continuous functions x with $x_0 = 0$) is the unique Borel probability on $\mathcal{C}$ such that

$$W(\{x : x_t - x_s \in B\}) = \frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \int_B \exp\left(\frac{-y^2}{2(t-s)}\right) dy$$

for $s < t$ and Borel $B \subset \mathbb{R}$, and such that disjoint increments $x_t - x_s$ of paths $x \in \mathcal{C}$ are independently distributed under W.

Take the hyperfinite time line $\overline{\mathbf{T}} = \mathbf{T} \cup \{1\}$, where $\mathbf{T}$ is as above and let $\mathcal{C}_N$ be the set of all polygonal paths $B(\mathsf{t})_{\mathsf{t}\in\mathbf{T}}$ filled in linearly between the time points $\mathsf{t} \in \mathbf{T}$, with $B(0) = 0$ and

$$B(\mathsf{t} + \Delta t) - B(\mathsf{t}) = \Delta B(\mathsf{t}) = \pm\sqrt{\Delta t}.$$

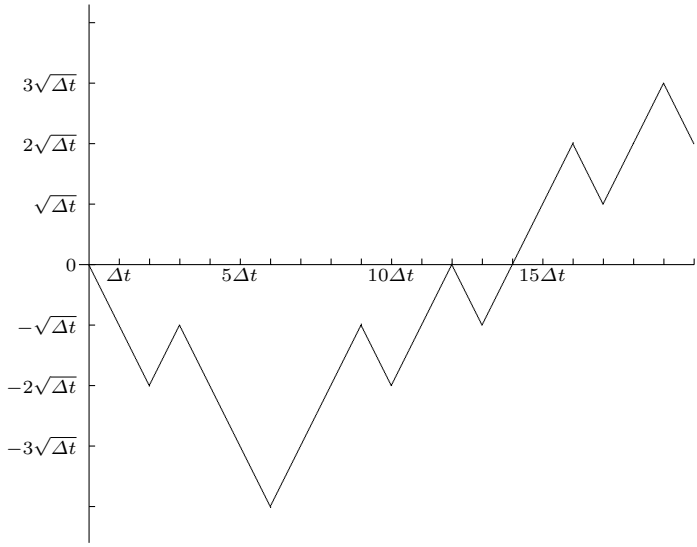
Let W_N =counting probability on $\mathcal{C}_N$, giving the internal probability space

$$(\mathcal{C}_N, \mathcal{A}_N, W_N)$$

where $\mathcal{A}_N = {}^*\mathcal{P}(\mathcal{C}_N)$. This gives the corresponding Loeb space

$$\boldsymbol{\Omega} = (\mathcal{C}_N, L(\mathcal{A}_N), P_N)$$

where $P_N = (W_N)_L$.



An infinitesimal random walk

Theorem 1.26 (Anderson) (a) *For a.a.*[19] $B \in \mathcal{C}_N$, *B is S-continuous, and gives a continuous path $b = {}^\circ B \in \mathcal{C}$.*

(b) *For Borel $D \subseteq \mathcal{C}$*

$$W(D) = P_N(\mathrm{st}^{-1}(D))$$

is Wiener measure.[20]

(c) *Writing $\Omega = \mathcal{C}_N$ and ω instead of B for a generic point in Ω, the process $b : [0,1] \times \Omega \to \mathbb{R}$ defined by*

$$b(t, \omega) = {}^\circ\omega(t)$$

is Brownian motion on the probability space $\boldsymbol{\Omega}$.

This is arguably the most intuitive of all the many constructions of Brownian motion/Wiener measure, and captures precisely the stochastic analyst's rule of thumb

$$\text{“}db^2 = dt\text{”},$$

since we really do have $\Delta B^2 = \Delta t$. Anderson [5] used it to give an elementary proof of Donsker's invariance principle, together with a pathwise construction of the Itô integral and an intuitive proof of Itô's Lemma. His construction opened the way for a large number of important applications in stochastic analysis and related fields, either directly or as an inspiration in more general situations. One of the first and most important of these, due to Keisler [61] is the idea of solving stochastic differential equations by means of *hyperfinite difference equations*. The paper [63] indicates some of the more recent developments in this area. We will discuss these ideas later (see section 1.5.3 below, and also Lecture 2), after we have outlined the basics of Loeb integration theory. First it is necessary to consider Loeb measurable *functions.*

1.3.4 Loeb measurable functions

Suppose we have a Loeb space $\boldsymbol{\Omega} = (\Omega, L(\mathcal{A}), \mu_L)$ constructed from the internal space $(\Omega, \mathcal{A}, \mu)$. A *Loeb measurable function* $f : \boldsymbol{\Omega} \to \overline{\mathbb{R}}$ is simply a function that is measurable in the conventional sense with respect to the Loeb algebra $L(\mathcal{A})$. That is, $f^{-1}(]-\infty, a]) \in L(\mathcal{A})$ for every real interval $[a, b]$.

There is of course another concept of measurable function, given by the transfer of the standard definition. A function $F : \boldsymbol{\Omega} \to {}^*\mathbb{R}$ is **measurable* if F is internal and $F^{-1}([\alpha, \beta]) \in \mathcal{A}$ for every hyperreal interval $[\alpha, \beta]$ (with $\alpha, \beta \in {}^*\mathbb{R}$). The fundamental connection between these two notions is as follows.

[19] with respect to the Loeb measure P_N of course.

[20] The standard part mapping here is the restriction to $\mathcal{C}_N$ of the mapping $\mathrm{st} : {}^*\mathcal{C} \to \mathcal{C}$ for the uniform topology – see section 1.2.4 above.

Theorem 1.27 *Let $f : \Omega \to \mathbb{R}$. Then the following are equivalent.*
(a) *f is Loeb measurable;*
(b) *there is a *measurable function $F : \Omega \to {}^*\mathbb{R}$ such that*

$$f(\omega) \approx F(\omega)$$

for almost all $\omega \in \Omega$ (with respect to the Loeb measure μ_L).[21]

For a proof see [29], [87], or the original paper of Anderson [6], who proved the result for measurable functions into a second-countable Hausdorff space. David Ross has extended this further to include *all* metric spaces [86].

Definition 1.28 A function F as given by Theorem 1.27 is called a *lifting* of f; that is, a *lifting* (with respect to μ_L) of a function $f : \Omega \to \mathbb{R}$ is an internal *measurable function $F : \Omega \to {}^*\mathbb{R}$ such that

$$f(\omega) \approx F(\omega)$$

for almost all $\omega \in \Omega$ (with respect to the Loeb measure μ_L).

A general lifting result that is very useful is Anderson's 'Luzin' theorem [6].

Theorem 1.29 *Let $(X, \mathcal{C}, \mu)$ be a complete Radon space and suppose that $f : X \to \mathbb{R}$ is measurable. Then $^*f$ is a lifting of f with respect to μ_L. That is,*

$$^*f(x) \approx f(^\circ x)$$

for $(^*\mu)_L$ almost all $x \in {}^*X$.

Remarks 1. The kind of lifting given by this theorem is known as a *two-legged* lifting, to distinguish it from the kind of lifting in Definition 1.28.

2. Anderson actually established this result for the situation where the range of f is any Hausdorff space with a countable base of open sets.

1.4 Loeb Integration Theory

Given a Loeb space $\boldsymbol{\Omega} = (\Omega, L(\mathcal{A}), \mu_L)$ and its originating internal space $(\Omega, \mathcal{A}, \mu)$, there are two integrals to consider. First, there is the *internal* integral

$$\int_\Omega F d\mu$$

[21] This result also holds for extended real valued functions $f : \Omega \to \overline{\mathbb{R}}$ provided we adopt the terminology $^\circ x = \pm\infty$ and hence $x \approx \pm\infty$ if $x \in {}^*\mathbb{R}$ is positive (resp. negative) infinite.

for any (internal) *integrable function $F : \Omega \to {}^*\mathbb{R}$. The value of this integral is a hyperreal that is given by the transfer of the construction of the integral on a standard space.

Secondly there is the classical Lebesgue integral

$$\int_\Omega f d\mu_L$$

defined in the usual way for a Loeb integrable function $f : \Omega \to \mathbb{R}$: the term *Loeb integrable function* f means simply that f is integrable (in the conventional sense) with respect to the Loeb measure μ_L on $\boldsymbol{\Omega}$.

Loeb integration theory gives the connection between these two integrals. Its importance stems from the fact that the internal integral $\int F d\mu$ may be quite simple (for example a hyperfinite sum) while a closely related Loeb integral can represent a general standard integral (such as a Lebesgue integral on the real line or a Wiener integral). Here are the details.

Theorem 1.30 *If F is a finitely bounded internal measurable function then*

$$\circ\int F d\mu = \int {}^\circ F d\mu_{\mathrm{L}}.$$

Corollary 1.31 *If F is a (finitely) bounded lifting of a Loeb measurable f, then*

$$\int f d\mu_{\mathrm{L}} = \circ\int F d\mu.$$

We cannot in general expect equality of ${}^\circ\int F d\mu$ and $\int {}^\circ F d\mu_{\mathrm{L}}$ since F may be large on a set of infinitesimal measure, as in the following example.

Example Consider $\Omega = {}^*[0,1]$ and define $F : {}^*[0,1] \to {}^*\mathbb{R}$ by

$$F(\tau) = \begin{cases} K \text{ for } \tau \leq \frac{1}{K} \\ 0 \quad \text{otherwise.} \end{cases}$$

Let Λ denote *Lebesgue measure. Then ${}^\circ F(\tau) = 0$ almost everywhere with respect to Λ_L, and hence $\int {}^\circ F d\Lambda_{\mathrm{L}} = 0$. But $\int F d\Lambda = 1$.

We always have

Theorem 1.32 *For any internal $\mathcal{A}$-measurable F with $F \geq 0$*

$$\int {}^\circ F d\mu_{\mathrm{L}} \leq \circ\int F d\mu,$$

where we allow the value ∞ on either side.

To obtain equality of $^\circ\int F d\mu$ and $\int {}^\circ F d\mu_{\mathrm{L}}$ it is necessary to have some condition on F akin to standard integrability — roughly, so that F is not too big on small sets. The following is the appropriate condition.

Definition 1.33 Let a function $F : \Omega \to {}^*\mathbb{R}$ be $\mathcal{A}$-measurable and internal and μ an internal finite measure. Then F is S-*integrable* if

(i) $\int_\Omega |F| d\mu$ is finite,

(ii) if $A \in \mathcal{A}$ and $\mu(A) \approx 0$, then $\int_A |F| d\mu \approx 0$.

Note If μ is not finite an extra condition has to be added:

(iii) if $A \in \mathcal{A}$ and $F \approx 0$ on A, then $\int_A |F| d\mu \approx 0$.

This is always satisfied for a finite measure μ. If $F \approx 0$ on A and $\mu(A) \neq 0$, then for any $0 < \varepsilon \in \mathbb{R}$ we have $|F| < \varepsilon$. Hence $\int_A |F| d\mu < \varepsilon \mu(A)$, which is enough since $\mu(A)$ is finite.

The function in the example above is not S-integrable because $A = [0, \frac{1}{K}]$ has $\Lambda(A) \approx 0$ but $\int_A F d\Lambda = 1$.

Note that F is S-integrable if and only if its positive and negative parts F^+ and F^- are S-integrable, and equivalently if $|F|$ is S-integrable.

The next result shows the importance of S-integrability.

Theorem 1.34 *Let $F : \Omega \to {}^*\mathbb{R}$ be $\mathcal{A}$-measurable with $F \geq 0$. Then the following conditions are equivalent:*

(a) *F is* S*-integrable,*

(b) *$^\circ F$ is Loeb integrable and*

$$^\circ\int F d\mu = \int {}^\circ F d\mu_{\mathrm{L}}.$$

The following is an equivalent formulation of S-integrability (the proof is left as an exercise).

Proposition 1.35 *An internal function F is* S*-integrable if and only if for all infinite K*

$$\int_{|F|>K} |F| d\mu \approx 0.$$

To complete the basic theory of Loeb integration we have:

Theorem 1.36 *Let $f : \Omega \to \mathbb{R}$ be Loeb measurable. Then f is μ_{L}-integrable if and only if it has an S-integrable lifting $F : \Omega \to {}^*\mathbb{R}$.*

Definition 1.37 We say that $F : \Omega \to {}^*\mathbb{R}$ is SL^p $(p > 0)$ if $|F|^p$ is S-integrable (so SL^1 means S-integrable).

Here is a very useful test for S-integrability isolated by Lindstrøm [68] and frequently applied in the case $p = 2$.

Theorem 1.38 *Suppose* $\mu(\Omega) < \infty$. *If* $F : \Omega \to {}^*\mathbb{R}$ *is internal,* $\mathcal{A}$*-measurable, and*

$$\int_\Omega |F|^p d\mu < \infty$$

for some $p > 1$, $p \in \mathbb{R}$, *then* F *is* S*-integrable.*

1.5 Elementary Applications

As a warm up for the more substantial applications of Loeb measures in later lectures, we present here a few simple illustrations of their power.

1.5.1 Lebesgue integration

Recall the hyperfinite time set $\mathbf{T}$ defined above (Definition 1.24), which carries the counting Loeb measure ν_L. For any function $f : [0,1] \to \mathbb{R}$ we may define a corresponding function $\hat{f} : \mathbf{T} \to {}^*\mathbb{R}$ by

$$\hat{f}(\mathsf{t}) = f({}^\circ\mathsf{t}).$$

The characterisation (or definition) of Lebesgue measure given by Theorem 1.25, combined with Theorem 1.27 yields immediately:

Theorem 1.39 *The following are equivalent:*

(a) f *is Lebesgue measurable;*

(b) $\hat{f}$ *is Loeb measurable (wrt* ν_L*);*

(c) *there is an internal function* $F : \mathbf{T} \to {}^*\mathbb{R}$ *(a lifting of* $\hat{f}$*) such that for a.a.* $\mathsf{t} \in \mathbf{T}$

$$f({}^\circ\mathsf{t}) = {}^\circ F(\mathsf{t})$$

The lifting F of $\hat{f}$ is a *two-legged lifting* in the sense described earlier.

Now apply Theorem 1.36 to give the following pleasant characterisation of the Lebesgue integral.

Theorem 1.40 *Suppose that* $f, \hat{f}$ *are as above. Then the following are equivalent:*

(a) f *is Lebesgue integrable*;

(b) $\hat{f}$ *is Loeb integrable*;

(c) *there is an* S*-integrable function* $F : \mathbf{T} \to {}^*\mathbb{R}$ *that is a lifting of* f *(and* $\hat{f}$*).*

If any of (a)–(c) *holds then*

$$\int_0^1 f d\lambda = \int_{\mathbf{T}} \hat{f} d\nu_L = {}^\circ \sum_{\mathsf{t} \in \mathbf{T}} F(\mathsf{t}) \Delta t,$$

the summation term $\sum_{\mathsf{t} \in \mathbf{T}} F(\mathsf{t}) \Delta t$ *being another way of writing* $\int_{\mathbf{T}} F d\nu$.

1.5.2 Peano's Existence Theorem

The above characterisation of the Lebesgue integral as a hyperfinite sum leads naturally to the method of *hyperfinite difference equations* for solving ODEs – an appealing technique pioneered by Keisler and extended to great effect especially for stochastic differential equations – see [61]. Here is an outline of a proof of Peano's fundamental existence theorem using this technique.

Theorem 1.41 (Peano) *Suppose that $f : [0,1] \times \mathbb{R} \to \mathbb{R}$ is bounded, measurable, and continuous in the second variable, and let $x_0 \in \mathbb{R}$. Then there is a solution to the differential equation*

$$\begin{aligned} dx(t) &= f(t, x(t))dt \\ x(0) &= x_0 \end{aligned} \tag{1.1}$$

(Of course, what is meant is really the corresponding integral *equation.)*

Proof Without any loss of generality we may assume that $x_0 = 0$ (otherwise consider the equation for $x(t) - x_0$). Suppose that $|f| \leq c$. An extension of the Lifting Theorem 1.27 is used (see below for details) to obtain an internal function $F : \mathbf{T} \times {}^*[-c, c] \to {}^*\mathbb{R}$ such that $|F| \leq c$ and for almost all $\mathsf{t} \in \mathbf{T}$

$$F(\mathsf{t}, X) \approx f({}^\circ\mathsf{t}, {}^\circ X) \tag{1.2}$$

for all $|X| \leq c$.

The *hyperfinite difference equation* corresponding to (1.1) is now

$$\Delta X(\mathsf{t}) = F(\mathsf{t}, X(\mathsf{t}))\Delta t,$$

where $\Delta X(\mathsf{t}) = X(\mathsf{t} + \Delta t) - X(\mathsf{t})$, together with the initial condition $X(0) = 0$. This is an internal equation for an internal function $X : \mathbf{T} \to {}^*\mathbb{R}$, with solution $X(\mathsf{t})$ defined recursively by

$$\begin{cases} X(0) = 0 \\ X(\mathsf{t} + \Delta t) = X(\mathsf{t}) + F(\mathsf{t}, X(\mathsf{t}))\Delta t. \end{cases}$$

Then X is S-continuous, and $|X(\mathsf{t})| \leq c\mathsf{t} \leq c$ for all $\mathsf{t} \in \mathbf{T}$. So we may define a continuous function $x : [0,1] \to \mathbb{R}$ by

$$x(t) = {}^\circ X(\mathsf{t})$$

for any $\mathsf{t} \approx t$. Clearly $|x(t)| \leq c$. To see that $x(t)$ is a solution, observe that, by (1.2) and the definition of x, for almost all $\mathsf{t} \in \mathbf{T}$

$$F(\mathsf{t}, X(\mathsf{t})) \approx f({}^\circ\mathsf{t}, {}^\circ X(\mathsf{t})) = f({}^\circ\mathsf{t}, x({}^\circ\mathsf{t}))$$

which means that the function $G(\mathsf{t}) = F(\mathsf{t}, X(\mathsf{t}))$ is a lifting of the function $g(t) = f(t, x(t))$. So, applying Theorem 1.40 to $g(t)$ and its lifting $G(\mathsf{t})$ we have (putting $t = {}^\circ\mathsf{t}$)

$$x(t) = {}^\circ X(\mathsf{t}) = {}^\circ\sum_{\mathsf{s}<\mathsf{t}} F(\mathsf{s}, X(\mathsf{s}))\Delta t$$
$$= \int_0^t f(s, x(s))ds$$

as required.

The lifting F above satisfying (1.2) is obtained as follows. Define the measurable function $\hat{f} : [0,1] \to C([-c,c])$ by

$$\hat{f}(t)(z) = f(t,z).$$

for $|z| \leq c$. From this we obtain (using Theorem 1.25) a Loeb measurable function $\check{f} : \mathbf{T} \to C([-c,c])$ (where $\mathbf{T}$ is the hyperfinite time line, endowed with the counting measure ν as above) by

$$\check{f}(\mathsf{t}) = \hat{f}({}^\circ\mathsf{t}).$$

for $\mathsf{t} \in \mathbf{T}$. Taking the uniform topology on $C([-c,c])$ and the extension of Theorem 1.27 to separable metric spaces, we obtain a lifting $\hat{F} : \mathbf{T} \to {}^*C([-c,c])$ such that for almost all $\mathsf{t} \in \mathbf{T}$ (with respect to ν_L)

$$\hat{F}(\mathsf{t}) \approx \check{f}(\mathsf{t}) = \hat{f}({}^\circ\mathsf{t})$$

(in the uniform topology) and $|\hat{F}| \leq c$. This means that for all such t

$$\hat{F}(\mathsf{t})(X) \approx \check{f}(\mathsf{t})({}^\circ X) = \hat{f}({}^\circ\mathsf{t})({}^\circ X) = f({}^\circ\mathsf{t}, {}^\circ X)$$

for all $|X| \leq c$. Now define $F : \mathbf{T} \times {}^*[-c,c] \to {}^*\mathbb{R}$ by

$$F(\mathsf{t}, X) = \hat{F}(\mathsf{t})(X).$$

Then $|F| \leq c$ and for almost all $\mathsf{t} \in \mathbf{T}$

$$F(\mathsf{t}, X) \approx f({}^\circ\mathsf{t}, {}^\circ X)$$

for all $|X| \leq c$, which is (1.2).

□

A slightly different Loeb measure approach to differential equations is to work with an infinitesimal *delayed equation*, and we illustrate this with an alternative proof of the Peano theorem.

Alternative Proof of Theorem 1.41 Let $\Delta = \Delta t = N^{-1}$ as above, and define an internal function $X : {}^*[-\Delta, 1] \to {}^*\mathbb{R}$ by

$$X(\tau) = x_0 \qquad \text{for } -\Delta \leq \tau \leq 0$$
$$X(\tau) = x_0 + \int_0^\tau {}^*f(\sigma, X(\sigma - \Delta))d\sigma \qquad \text{for } 0 \leq \tau \leq 1$$

Note that $X(\tau)$ is defined recursively on $[k\Delta, (k+1)\Delta]$ for $0 \leq k < N$.

Since f and hence ${}^*f$ is bounded, X is S-continuous and we can define a standard function $x : [0,1] \to \mathbb{R}$ by

$$x(t) = {}^\circ X(t) = {}^\circ X(\tau)$$

for any $\tau \approx t$. We claim that $x(t)$ is a solution to equation (1.1).

Let $\Lambda = {}^*\lambda = {}^*$Lebesgue measure. Using the extension of Anderson's Luzin Theorem 1.29, mentioned above, and considering the function $\hat{f} : [0,1] \to C(\mathbb{R})$ defined by $\hat{f}(t)(z) = f(t,z)$ we have that for almost all τ (with respect to Λ_L)

$${}^*f(\tau, y) \approx f({}^\circ\tau, {}^\circ y) \qquad \text{for all finite } y \in {}^*\mathbb{R}.$$

Hence, for almost all $\tau \in {}^*[0,1]$

$${}^*f(\tau, X(\tau - \Delta)) \approx f({}^\circ\tau, {}^\circ X(\tau - \Delta)) = f({}^\circ\tau, x({}^\circ\tau))$$

since ${}^\circ(\tau - \Delta) = {}^\circ\tau$. Now this means that $G(\tau) = {}^*f(\tau, X(\tau - \Delta))$ is a bounded lifting of $g(\tau) = f({}^\circ\tau, x({}^\circ\tau))$ and so for any $t \in [0,1]$

$$x(t) = {}^\circ X(t) = x_0 + {}^\circ\int_0^t G(\tau)d\tau = x_0 + \int_0^t g(\tau)d_L\tau$$

where $d_L\tau$ denotes integration with respect to Λ_L. Since $\Lambda_L \circ \mathrm{st}^{-1}$ is Lebesgue measure, we have

$$\int_0^t g(\tau)d_L\tau = \int_0^t f({}^\circ\tau, x({}^\circ\tau))d_L\tau = \int_0^t f(t, x(t))dt$$

which shows that $x(t)$ is a solution to equation (1.1). □

Loeb Differential Equations

The existence of the Loeb-Lebesgue measure ${}^*\lambda_L$ on ${}^*\mathbb{R}$ makes it possible (and natural) to formulate and solve *Loeb differential equations* for the "rich" time line ${}^*\mathbb{R}$. By this we mean *integral* equations of the following kind:

$$x(\tau) = x_0 + \int_0^\tau f(\sigma, x(\sigma))d_L\sigma$$

where $f : {}^*[0,1] \times \mathbb{R} \to \mathbb{R}$ is Loeb measurable in τ and continuous in the second variable. The solution $x(\tau)$ will be S-continuous *and* real valued, so it will really be a continuous function. Such equations occur in the study of optimal control theory, where it is natural to consider Loeb measurable controls. In particular, it can be shown [55] that a general optimal control problem will always have an optimal Loeb control, even when there is no optimal Lebesgue control. There are close connections here with Young measures: this was shown in the context of control theory in [31], and is discussed in greater detail in the forthcoming paper [45].

1.5.3 Itô integration and stochastic differential equations

The hyperfinite difference approach has been used to great effect in the solution of Itô stochastic differential equations (SDEs), based on Anderson's hyperfinite random walk construction of Brownian motion and the Itô integral [5].

Without going into details, the *Itô integral* gives a meaning to the expression

$$I(t,\omega) = \int_0^t f(s,\omega)db(s,\omega)$$

where $b(t,\omega)$ is Brownian motion and $f(t,\omega)$ is a certain kind of random function (an *adapted function).* The Itô integral $I(t,\omega)$ is a continuous stochastic process, and is defined as an L^2-limit of simpler random processes. In the standard theory a pathwise (that is, ω-wise) definition of I is not possible. Nevertheless Anderson [5] showed how to represent the Itô integral pathwise as a hyperfinite sum, in a direct generalisation of the above representation of the Lebesgue integral.

First recall Anderson's Brownian motion $b(t,\omega)$ constructed earlier on the Loeb space $\boldsymbol{\Omega} = (\mathcal{C}_N, L(\mathcal{A}_N), P_N)$. This came from the canonical internal random walk $B(\mathsf{t},\omega)$ defined on $\Omega = \mathcal{C}_N$ by

$$B(\mathsf{t},\omega) = \omega(\mathsf{t})$$

Thus $\Delta B(\mathsf{t},\omega) = B(\mathsf{t}+\Delta t,\omega) - B(\mathsf{t},\omega) = \pm\sqrt{\Delta t}$.

A generalisation of Theorem 1.27 gives:

Theorem 1.42 *Let $f(t,\omega)$ be an adapted function. Then there is a nonanticipating*[22] *lifting $F : \mathbf{T}\times\Omega \to {}^*\mathbb{R}$ of f such that*

$$f({}^\circ\mathsf{t},\omega) \approx F(\mathsf{t},\omega)$$

for almost all $(\omega,\mathsf{t}) \in \mathbf{T}\times\Omega$.

Anderson [5] proved the following stochastic generalisation of Theorem 1.40.

Theorem 1.43 (Anderson) *Let F be nonanticipating lifting of an adapted function f as above, and define an internal hyperfinite stochastic integral $G : \mathbf{T}\times\Omega \to {}^*\mathbb{R}$ by*

$$G(\mathsf{t},\omega) = \sum_{\mathsf{s}<\mathsf{t}} F(\mathsf{s},\omega)\Delta B(\mathsf{s},\omega).$$

Then G is a nonanticipating lifting of the stochastic integral $I(t,\omega) = \int f db$ defined above. That is, for almost all ω, the function $G(\mathsf{t},\omega)$ is S-continuous and

[22] this means that $F(\mathsf{t},\omega)$ depends only on the values $\omega(\mathsf{s})$ for $\mathsf{s} \le \mathsf{t}$

$$I({}^{\circ}\mathtt{t}, \omega) \approx G(\mathtt{t}, \omega)$$

for all $\mathtt{t}$.

In [61] Keisler pioneered the use of Anderson's representation of the Itô integral in the solution of stochastic differential equations (SDEs), generalising the technique described above to prove the Peano Existence Theorem for ODEs. Subsequently this technique has been developed by many authors, both in solving SDEs and in applications such as optimal control theory [22] and mathematical finance theory ([35] for example). The delay approach is also appropriate for certain SDEs – see [21]. Loeb space methods for SDEs have been extended to equations involving general stochastic integrals against martingales and semimartingales, beginning with Hoover & Perkins [57] and Lindstrøm [68].

For partial differential equations (PDEs) – or, more generally, infinite dimensional differential equations, in addition to the above approaches there are new possibilities for constructing solutions using hyperfinite *dimensional* representation of the objects concerned (and not necessarily using hyperfinite representation of time). The book [13] develops this idea in some detail for the Navier–Stokes equations, which are formulated as a differential equation in a certain separable Hilbert space $\mathbf{H}$. We will discuss this in greater detail in the next lecture – which includes some developments since the publication of [13].

Measure valued equations on an infinite dimensional space can also be treated successfully using hyperfinite dimensional representation, together with the idea of *nonstandard densities* – and this is also touched upon in the next lecture.