# NONSTANDARD MEASURE THEORY AND ITS APPLICATIONS

NIGEL J. CUTLAND


| | | Page |
|---|---|---|
| *INTRODUCTION AND PRELIMINARIES* | | 530 |
| §1 | NONSTANDARD MATHEMATICS | 533 |
| | (1) *The hyperreals* $^*\mathbb{R}$ *and nonstandard analysis* | 533 |
| | (2) *The nonstandard universe* $^*V(\mathbb{R})$ | 536 |
| | (3) *Some basic nonstandard tools* | 540 |
| | (4) *Nonstandard topology.* | 543 |
| | (5) *Metric spaces* | 544 |
| | (6) *Measure spaces.* | 544 |
| §2 | EARLY NONSTANDARD MEASURE THEORY | 545 |
| | (1) *Lebesgue measure* | 545 |
| | (2) *Finitely additive measures* | 547 |
| §3 | LOEB MEASURES | 549 |
| | (1) *Construction of Loeb measures* | 549 |
| | (2) *Loeb measurable functions and integration* | 552 |
| | (3) *Hyperfinite Loeb spaces* | 556 |
| | (4) *Products of Loeb spaces* | 557 |
| §4 | REPRESENTATIONS OF MEASURES BY LOEB MEASURES | 558 |
| | (1) *Lebesgue measure* | 558 |
| | (2) *Measures on Hausdorff spaces* | 560 |
| | (3) *Representation of general measures* | 561 |
| §5 | CONSTRUCTION OF MEASURES | 562 |
| | (1) *Extension of measures.* | 562 |
| | (2) *Weak limits of measures* | 564 |
| §6 | PROBABILITY AND STOCHASTIC PROCESSES | 565 |
| | (1) *Universality of hyperfinite Loeb spaces* | 566 |
| | (2) *Construction of a Poisson process* | 566 |
| | (3) *Anderson's Brownian motion.* | 568 |
| | (4) *White noise and Lévy Brownian motion* | 571 |


Received 3 May, 1983.

*Bull. London Math. Soc.*, 15 (1983), 529–589

§7 STOCHASTIC ANALYSIS .. .. .. .. .. .. .. .. 571
(1) *Itô integration* .. .. .. .. .. .. .. .. 571
(2) *Stochastic differential equations* .. .. .. .. .. .. 575
(3) *Stochastic integration* .. .. .. .. .. .. .. .. 577
(4) *Other topics in stochastic analysis* .. .. .. .. .. 579

§8 ADAPTED DISTRIBUTIONS .. .. .. .. .. .. .. .. 579

§9 FURTHER APPLICATIONS .. .. .. .. .. .. .. .. 581
(1) *Self-avoiding random walks* .. .. .. .. .. .. 581
(2) *Optimal control theory* .. .. .. .. .. .. .. 582
(3) *Infinite particle systems* .. .. .. .. .. .. .. 583
(4) *Mathematical economics* .. .. .. .. .. .. .. 584
(5) *The Boltzmann equation* .. .. .. .. .. .. .. 584
(6) *The ergodic theorem* .. .. .. .. .. .. .. 585

APPENDIX: CONSTRUCTION OF $^*V(\mathbb{R})$ .. .. .. .. .. .. .. 585

*REFERENCES* .. .. .. .. .. .. .. .. .. .. 587

## *Introduction*

Nonstandard measure theory has been used in recent years to establish new *standard* results in a variety of areas—for example the extension of measures, potential theory, Brownian local time, stochastic differential equations, optimal control theory, adapted probability distributions, and the modelling of infinite particle systems, exchange economies and polymer chains. Most of these applications stem from P. Loeb's discovery [**Lo$_4$**] of a way to construct new rich *standard* measure spaces from nonstandard measure spaces. These have become known as *Loeb spaces.* Loeb's construction is, with hindsight, quite simple. Nonstandard methods also provide fresh and often very intuitive and illuminating ways to approach established standard theory: a beautiful example of this (which has had significant applications) is Anderson's construction of Brownian motion from a random walk having genuinely infinitesimal steps.

Important among Loeb spaces are those that are *hyperfinite.* These are infinite standard measure spaces that are finite from the nonstandard viewpoint, and hence behave formally like finite spaces. Lebesgue measure and integration on $[0, 1]$, for example, can be defined from the simple counting probability measure on a hyperfinite set $T = \{0, \Delta t, 2\Delta t, ..., 1\}$, where $\Delta t$ is a positive infinitesimal; integration on $T$ is just summation. Anderson's Brownian motion lives on a hyperfinite Loeb space given by a counting probability measure. In a sense that can be made precise, hyperfinite Loeb spaces are *universal*, and this accounts for their wide range of applicability. They have been shown to be fruitful in the modelling of large finite situations.

Our aim in this paper is to give a comprehensive survey of the achievements and methods of nonstandard measure theory and its standard applications, beginning with a discussion of the state of the art before the advent of Loeb measure.

Subsequently we restrict attention almost exclusively to Loeb measures and their use in a wide variety of fields, as can be seen from the table of contents above.

The discovery of nonstandard analysis was made in 1960 by Abraham Robinson, who used ideas from logic to construct an enrichment $^*\mathbb{R}$ of the reals $\mathbb{R}$ containing infinitesimals and infinite elements, and inheriting all the properties of $\mathbb{R}$ (suitably defined). The book **[Rb]** describes this and also Robinson's extension of nonstandard techniques to other areas of mathematics. The basic idea is to embed every mathematical structure $\mathfrak{A}$ in an extension $^*\mathfrak{A}$ which contains 'ideal' elements. (This is somewhat akin to the extension of the rationals to the reals, or the completion of any metric space, by the inclusion of 'ideal' elements.) A suitable extension $^*\mathfrak{A}$ can be obtained algebraically by taking $^*\mathfrak{A} = \mathfrak{A}^I/\mathcal{U}$, an *ultrapower*, where $\mathcal{U}$ is an ultrafilter on the subsets of the index set $I$; the ultrapower is just a quotient of the cartesian product $\mathfrak{A}^I$. Thus nonstandard methods could in principle be routinely eliminated (along with their beauty and clarity) in favour of ultrapower (or ultraproduct) techniques; however this would be extremely tedious.

Nonstandard methods provide extra tools and concepts for the working mathematician, to be used alongside and in conjunction with, not necessarily instead of, other tools. Many of the results referred to in the opening paragraph (and discussed later) do not have any alternative proof. As far as we know, Loeb spaces are the only spaces having the strong universality properties discussed in §8. Thus nonstandard tools in mathematics are *not* (as unfortunately suggested by the term *nonstandard*) merely an alternative to standard methods, or just a curiosity invented by eccentric logicians. The evidence of the developments surveyed in this paper suggests that nonstandard ideas provide fresh power and insight, and should not be overlooked, at least in measure theory and related fields. Perhaps the term nonstandard should be scrapped in favour of the term *hyperstandard* or *ultrastandard*.

In this paper we assume no previous knowledge of nonstandard methods. An axiomatic approach to nonstandard mathematics now available is designed so that the mathematician interested in using this tool need not first become initiated in the mysteries of mathematical logic. We have included a fairly lengthy section (§1) developing the basics of nonstandard mathematics using this approach; for the sake of completeness and to eliminate any air of mystery, we give in the appendix an algebraic construction of a nonstandard universe satisfying the axioms in §1. Nonstandard methods are demonstrated in action both in §1 and in later sections, where at times we have given detailed proofs. Space forbids a totally comprehensive and leisurely introduction to nonstandard ideas, and the books **[D, Ke$_2$, SL, Rb]** may be consulted for further information.

The mathematical assumptions in this paper are quite minimal, consisting mainly of familiarity with the rudiments of conventional measure theory and integration. When discussing specialised topics we have either explained some of the basic ideas involved or else given only a sketchy treatment. Our aim is to appeal to a wide audience, and yet provide sufficient meat to attract the interest of experts.

The depth of exposition varies from topic to topic through the survey. In developing the basic Loeb theory (§3), for example, we have given full proofs, because the later sections rest on this and all existing textbooks stop short of Loeb measure. (This will be remedied when the books **[AFHL]** and **[BS]** eventually appear.) To take another example, Anderson's Brownian motion is presented in some detail as an illustration of the power and illumination provided by nonstandard techniques. (It is also a pleasant way to discover what Brownian motion is and how to construct it.) Some topics, on the other hand, are touched on quite briefly and in less than full generality. Naturally the selections made have been influenced not only by the desire to illuminate rather than swamp the reader, but also by the limitations of space and by the author's own interests.

Much of the work reported in this survey is carefully expounded in one or both of the forthcoming books [**AFHL**, **BS**], often with greater generality than in the source papers, and with many helpful examples and elaborations.

Thanks are due to the many people from whom nonstandard measure theory has been learned; special thanks are offered to Jerry Keisler, who introduced me to this subject in his graduate lectures; I have greatly appreciated helpful and stimulating contact also with Bob Anderson, Jens Erik Fenstad, Ward Henson, Doug Hoover, Al Hurd, Steve Kosciuk, Tom Lindstrøm, Peter Loeb, Ed Perkins, Keith Stroyan and Frank Wattenberg.

### *Mathematical preliminaries*

It will be helpful to note now some of the basic terminology and notation that we shall use; this is intended to be entirely conventional.

The symbol $\bar{\mathbb{R}}$ denotes the extended real numbers $\mathbb{R} \cup \{-\infty, \infty\}$, in the usual sense, and $\bar{\mathbb{R}}^+$ denotes $\mathbb{R}^+ \cup \{\infty\}$, where $\mathbb{R}^+ = [0, \infty)$. The set of natural numbers $\mathbb{N}$ is $\{1, 2, 3, \ldots\}$ and $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$.

For any set $A$, the power set of $A$ (the set of all subsets of $A$) is written $\mathscr{P}(A)$; for sets $A, B$ the set theoretic difference is written $A \setminus B$, and $A \,\Delta\, B$ denotes the symmetric difference $(A \setminus B) \cup (B \setminus A)$. The complement of $A$ in a given context is written $A^c$.

If $f$ is a function, the restriction of $f$ to the set $A$ is written $f \upharpoonright A$. If we say *function* without further specification we mean a real function. For a topological space $X$ the set of continuous functions from $X$ to $\mathbb{R}$ is denoted $C(X)$. If $f : X \to Y$ is a function we use interchangeably the notation $f_x$ and $f(x)$ for the value of $f$ for $x$ in $X$.

An *algebra* of subsets of a set $X$ is a subset of $\mathscr{P}(X)$ that is closed under complements and unions (hence differences and intersections). A *$\sigma$-algebra* is an algebra that is closed under countable unions (and intersections). A *measurable space* is a pair $(X, \mathscr{F})$ where $\mathscr{F}$ is a $\sigma$-algebra on $X$. For any collection $\mathscr{D} \subseteq \mathscr{P}(X)$, the smallest $\sigma$-algebra containing $\mathscr{D}$ is denoted by $\sigma(\mathscr{D})$. For two collections $\mathscr{D}, \mathscr{E}$ the $\sigma$-algebra $\sigma(\mathscr{D} \cup \mathscr{E})$ is denoted by $\mathscr{D} \vee \mathscr{E}$.

The term *measure space* without further qualification means a triple $\mathbf{X} = (X, \mathscr{F}, \mu)$ where $\mathscr{F}$ is a $\sigma$-algebra on $X$ and $\mu : \mathscr{F} \to \bar{\mathbb{R}}^+$ is countably additive. A *finitely additive measure space* is a triple $\mathbf{X}$ for which $\mathscr{F}$ is an algebra and $\mu$ is only finitely additive. We abbreviate finitely additive by f.a. A measure (or a f.a. measure) $\mu$ is *finite* if $\mu(X) < \infty$; it is a *probability measure* (or a *f.a. probability measure*) if $\mu(X) = 1$. A *signed measure* (or *signed f.a. measure*) is one taking positive and negative values (but not both $+\infty$ and $-\infty$).

For a fixed measure space $\mathbf{X} = (X, \mathscr{F}, \mu)$ a *null set* is a subset of a set in $\mathscr{F}$ of measure zero; denoting the null sets by $\mathscr{N}$, the *completion* of $\mathscr{F}$ (with respect to $\mu$) is $\mathscr{F} \vee \mathscr{N}$; to obtain the completion of $\mathbf{X}$ the measure $\mu$ is extended in the obvious way. We use the conventional abbreviations a.a. (*almost all*) and a.s. (*almost surely*) to describe phenomena that hold for all $x \in X$ except for a null set.

We use the conventional notation $\mathscr{L}^1(\mathbf{X})$ for the integrable functions on a measure space $\mathbf{X}$, and $\mathscr{L}^\infty(\mathbf{X})$ denotes those that are essentially bounded. The distinction between $\mathscr{L}^1$ (etc.) and the Lebesgue spaces $L^1$ (etc.) is blurred as is customary; however we shall use $\mathscr{L}^1$ (etc.) rather than $L^1$ (etc.) to avoid possible confusion with a later use of the symbol $L$ in connection with Loeb spaces, which has become conventional. (Strictly there would be no confusion as $L$ never has superscripts in the Loeb context, but it is a little too close for comfort.)

## §1 Nonstandard Mathematics

We begin this section with a description of the hyperreals $^*\mathbb{R}$ and some fundamental nonstandard notions involving them, followed by a construction of $^*\mathbb{R}$ and a sample of nonstandard real analysis. In §1(2) we give an axiomatisation of the nonstandard (or hyperreal) universe $^*V(\mathbb{R})$ that is needed to do more general nonstandard mathematics. The rest of this section discusses briefly those nonstandard ideas and methods that will be needed later. For fuller treatments of the basics of nonstandard mathematics the books [**D**] or [**SL**] should be consulted.

### §1(1) *The hyperreals* $^*\mathbb{R}$ *and nonstandard analysis*

Nonstandard real analysis works with an ordered field $^*\underset{\sim}{\mathbb{R}} = (^*\mathbb{R}, ^*+, ^*\times, ^*<, 0, 1)$ that properly extends the real ordered field and contains both infinite and infinitesimal members. The[(1)] structure $^*\underset{\sim}{\mathbb{R}}$ is constructed (or accepted axiomatically) so that it inherits "all properties" of the standard reals. At the same time each function $f : \mathbb{R}^n \to \mathbb{R}$ is extended to a function $^*f : (^*\mathbb{R})^n \to {}^*\mathbb{R}$, and each set $S \subseteq \mathbb{R}^n$ has an extension $^*S \subseteq (^*\mathbb{R})^n$ such that the properties of $f$ and $S$ are inherited by $^*f$ and $^*S$. Thus, for example, taking $f = |\cdot|$,

$$^*|x| = \begin{cases} x & \text{if } x \,^*\!\geqslant 0 \\ -x & \text{if } x \,^*\!< 0 \end{cases}$$

because this is a property of $|\cdot|$. It is customary to drop the prefix $^*$ from $^*|\cdot|$ and similarly for other basic functions and relations such as $^*+$, $^*\times$, $^*<$, etc.

The following definitions and results are fundamental (and make sense for any ordered field extending $\mathbb{R}$).

1.1. Definition. Let $x, y \in {}^*\mathbb{R}$;

(a) $x$ is *finite* if $|x| < n$ for some $n \in \mathbb{N}$;
$x$ is *infinite* if $|x| > n$ for all $n \in \mathbb{N}$;

(b) $x$ is *infinitesimal* if $|x| < \dfrac{1}{n}$ for all $n \in \mathbb{N}$; this is written $x \approx 0$;

(c) $x$ is *infinitely close* to $y$ if $|x-y| \approx 0$; this is written $x \approx y$.

[(1)] Unless certain somewhat unnatural specifications are made, there is not a *unique* structure $^*\underset{\sim}{\mathbb{R}}$ (even to within isomorphism); normally, however, one such $^*\underset{\sim}{\mathbb{R}}$ is fixed and is regarded as *the* hyperreals. This is akin to talking about *the* reals, which are actually only determined uniquely once a universe of set theory and a meaning for 'all sequences of rationals' has been fixed.

1.2. STANDARD PART THEOREM. *Let* $x \in {}^*\mathbb{R}$ *with* $x$ *finite; then there is a unique* $r \in \mathbb{R}$ *with* $x \approx r$; $r$ *is the* standard part *of* $x$ *and is denoted* $^\circ x$ *or* $\mathrm{st}(x)$.

*Proof.* Take $r = \sup\{s \in \mathbb{R} : s \leqslant x\}$, which exists because $x$ is finite.

1.3. COROLLARY. *There are infinite elements and non-zero infinitesimal elements in* ${}^*\mathbb{R}$.

*Proof.* Take $x \in {}^*\mathbb{R} \setminus \mathbb{R}$; if $x$ is finite then $0 \neq x - {}^\circ x \approx 0$, and $(x - {}^\circ x)^{-1}$ is infinite; if $x$ is infinite then $0 \neq x^{-1} \approx 0$.

It is convenient to define $^\circ x$ ( $= \mathrm{st}(x)$) for infinite $x$ by

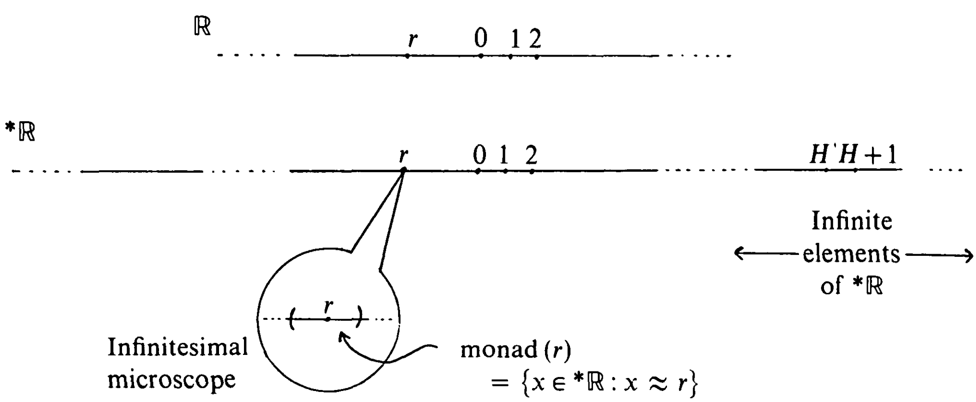
$$^\circ x = \begin{cases} \infty & \text{if } x > 0 \\ -\infty & \text{if } x < 0 \end{cases}$$

so that $^\circ = \mathrm{st} : {}^*\mathbb{R} \to \overline{\mathbb{R}}$.

1.4. DEFINITION. For $r \in \mathbb{R}$ the *monad* of $r$ is the set $\mathrm{monad}(r) = \{x \in {}^*\mathbb{R} : x \approx r\}$.

The set ${}^*\mathbb{N}$ is an important subset of ${}^*\mathbb{R}$; like $\mathbb{N}$ it is discrete. If $x \in {}^*\mathbb{R}$ is infinite there is a unique $H \in {}^*\mathbb{N}$ with $H \leqslant x < H+1$ (because this is a property of $\mathbb{R}$ and $\mathbb{N}$); so ${}^*\mathbb{N}$ has infinite members.

The hyperreals should be envisaged as follows.




Here briefly is one way to construct ${}^*\mathbb{R}$. Let $I$ be a countably infinite set and let $\mathscr{U}$ be an ultrafilter on $I$ that extends the filter generated by the cofinite subsets of $I$. Define ${}^*\mathbb{R}$ to be the quotient of $\mathbb{R}^I$ under the equivalence $\sim$ given by

$$x \sim y \Leftrightarrow \{i : x_i = y_i\} \in \mathscr{U}.$$

Write $x/\mathscr{U}$ for the equivalence class of $x$ and write ${}^*\mathbb{R} = \mathbb{R}^I/\mathscr{U}$.

For a function $f : \mathbb{R} \to \mathbb{R}$ define ${}^*f$ by ${}^*f(x/\mathscr{U}) = (f(x_i))_{i \in I}/\mathscr{U}$ (checking that this is well defined); and for $S \subseteq \mathbb{R}$ define ${}^*S \subseteq {}^*\mathbb{R}$ by

$$x/\mathscr{U} \in {}^*S \Leftrightarrow \{i : x_i \in S\} \in \mathscr{U}.$$

If $f$ or $S$ is $n$-ary, define ${}^*f$ and ${}^*S$ similarly.

Finally embed $\mathbb{R}$ in $^*\mathbb{R}$ by identifying $r \in \mathbb{R}$ with $\mathbf{r}/\mathscr{U}$ where $\mathbf{r} \in \mathbb{R}^I$ is constant with value $r$.

The precise sense in which $^*\underset{\sim}{\mathbb{R}}$ inherits all the properties of $\underset{\sim}{\mathbb{R}}$ is the following.

1.5. Transfer Principle. *Let $\phi(f_1, \ldots, f_k, S_1, \ldots, S_m, r_1, \ldots, r_n)$ be a statement involving $+$, $\times$, $<$, $0$, $1$, the real functions $f_1, \ldots, f_k$, the sets (or relations) $S_1, \ldots, S_m$, and the real numbers $r_1, \ldots, r_n$, such that $\phi$ can be expressed using the notions $=$, $\in$, &, or, not, $\Rightarrow$, $\forall x$, $\exists y$, ..., where $x, y, \ldots$ are real variables. Then the following are equivalent:*

(a) *$\phi(f_1, \ldots, f_k, S_1, \ldots, S_m, r_1, \ldots, r_n)$ is true in $\underset{\sim}{\mathbb{R}}$;*

(b) *$\phi(^*f_1, \ldots, {}^*f_k, {}^*S_1, \ldots, {}^*S_m, r_1, \ldots, r_n)$ is true in $^*\mathbb{R}$ (with the variables $x$, $y$ in $\exists x$, $\forall y$ ranging over $^*\mathbb{R}$).*

The transfer principle is a *theorem* if working with a specifically constructed $^*\mathbb{R}$; it is an *axiom* if working axiomatically.

The use of transfer in nonstandard real analysis is illustrated by the next theorem.

1.6. Theorem. *Let $f$ be a real function. Then $f$ is continuous at $c \in \mathbb{R}$ if and only if $^*f(x) \approx {}^*f(c)$ whenever $x \approx c$.*

*Proof.* If $f$ is continuous at $c$, then for each $n \in \mathbb{N}$ there is $\delta_n > 0$ such that in $\mathbb{R}$

$$\forall x\big(|x-c| < \delta_n \Rightarrow |f(x)-f(c)| < 1/n\big).$$

The transfer principle tells us that in $^*\mathbb{R}$

$$\forall x\big(|x-c| < \delta_n \Rightarrow |^*f(x)-{}^*f(c)| < 1/n\big).$$

In particular, if $x \approx c$, then $|x-c| < \delta_n$ for each $n \in N$ and so $|^*f(x) - {}^*f(c)| < 1/n$, for each $n$. Hence $^*f(x) \approx {}^*f(c)$.

Conversely, suppose that $^*f(x) \approx {}^*f(c)$ whenever $x \approx c$. Take any real $\varepsilon > 0$ and let $\Delta > 0$, $\Delta \approx 0$ in $^*\mathbb{R}$. In $^*\mathbb{R}$ we have

$$\forall x\big(|x-c| < \Delta \Rightarrow |^*f(x)-{}^*f(c)| < \varepsilon\big)$$

and hence, in $^*\mathbb{R}$

$$\exists\delta\forall x\big(|x-c| < \delta \Rightarrow |^*f(x)-{}^*f(c)| < \varepsilon\big).$$

By transfer the same holds in $\mathbb{R}$ with $f$ instead of $^*f$.

The above intuitive characterisation of continuity gives an easy proof of the intermediate value theorem. Suppose that $f$ is continuous on $[0, 1]$ and $f(0) < 0 < f(1)$. Take infinite $H \in {}^*\mathbb{N}$; then by transfer of a suitable statement about $\mathbb{R}$, $\mathbb{N}$ and $f$ there is $k \in {}^*\mathbb{N}$ such that

$$0 < k \leqslant H \quad \text{and} \quad {}^*f(k-1/H) < 0 \leqslant {}^*f(k/H).$$

Let $c = {}^\circ(k-1/H) = {}^\circ(k/H)$; then $f(c) \approx {}^*f(k-1/H) \approx {}^*f(k/H)$, so $f(c) = 0$.

A real sequence $s = (s_n)_{n\in\mathbb{N}}$, being a function $s:\mathbb{N}\to\mathbb{R}$, has an extension $^*s : {}^*\mathbb{N} \to {}^*\mathbb{R}$ which we denote by $(s_n)_{n\in{}^*\mathbb{N}}$. The following theorem is basic; we leave the proof (using ideas similar to those in Theorem 1.6) as a limbering up exercise.

1.7. THEOREM. *Let $(s_n)$ be a real sequence, and let $l \in \overline{\mathbb{R}}$.*

(a) *The limit points of $(s_n)$ are the points $^\circ s_H$ for infinite $H$.*

(b) *$s_n \to l$ as $n \to \infty$ if and only if $^\circ s_H = l$ for all infinite $H \in {}^*\mathbb{N}$.*

Clearly much of the calculus of continuity and convergence follows easily from the above characterisations. As a final illustration of basic nonstandard real analysis we mention Riemann integration.

1.8. THEOREM. *Let $f$ be a continuous on $[0,1]$ and let $H\in{}^*\mathbb{N}$ be infinite. Let $\Delta t = H^{-1}$; then the sum*

$$^\circ\left(\sum_{k=1}^{H} {}^*f(k/H)\Delta t\right)$$

*is independent of the choice of $H$, and equals $\int_0^1 f$.*

(The expression above means $\sigma_H$, where $\sigma_n$ is defined for finite $n$ by $\sigma_n = \sum_{k=1}^{n} f(k/n)n^{-1}$.)

One of the themes of later sections is the extent to which integration in general measure spaces can be similarly characterised.

A simple axiom system for $^*\mathbb{R}$ used in teaching undergraduate calculus has been developed by Keisler [**Ke**$_2$]; his axioms do not mention the transfer principle (although it is implied by them). A variant on Keisler's axioms was given by Tall [**T**].

## §1(2) *The nonstandard universe* $^*V(\mathbb{R})$

A standard universe sufficient for most general mathematics is the superstructure $V(\mathbb{R})$, where $V(S)$ is defined for any set $S$ by

$$V_0(S) = S;\quad V_{n+1}(S) = V_n(S) \cup \mathscr{P}(V_n(S));\quad V(S) = \bigcup_{n\in\mathbb{N}} V_n(S).$$

The set $S$ is to be regarded as a set of *atoms* (or *urelements*); i.e. if $x \in S$ then $x$ has no members. Should $V(\mathbb{R})$ not be big enough we could replace $\mathbb{R}$ by some set $S \supseteq \mathbb{R}$ to ensure that $V(S)$ contains all mathematical objects of interest. (We are assuming that all mathematical objects can be defined ultimately as sets.)

The[(2)] *nonstandard universe* $^*V(\mathbb{R})$ has the following informal description.

(i) *Every* standard object $A \in V(\mathbb{R})$ has a nonstandard counterpart $^*A$, which is an object in $^*V(\mathbb{R})$ of the same kind as $A$. For example, if $A$ is a set, then so is $^*A$;

[(2)] The comments made in the previous footnote regarding the uniqueness of $^*\mathbb{R}$ apply here to the question of the uniqueness of $^*V(\mathbb{R})$.

if $A$ is a set of functions then so is $^*A$; if $A$ is a triple $(X, \mathcal{F}, \mu)$ then $^*A$ is a triple (and in fact $^*A = (^*X, ^*\mathcal{F}, ^*\mu)$).

(ii) The objects $^*A$ possess the 'same properties' (intrinsically *and* in relation to each other) as their standard counterparts $A$. Thus if $A$ is a set, for any $x, y$ we have $x \in A \Leftrightarrow {}^*x \in {}^*A$ and $x = y \Leftrightarrow {}^*x = {}^*y$. If $f$ is a function and $f(x) = y$ then $^*f$ is a function and $^*f(^*x) = {}^*y = {}^*(f(x))$.

Taking $A = C[0, 1]$ for example, since $A$ is a collection of functions from $[0, 1]$ to $\mathbb{R}$, then $^*A$ is a collection of functions from $^*[0, 1]$ to $^*\mathbb{R}$. If we have $B = (X, \mathcal{F}, \mu)$ and $\mathcal{F}$ is a collection of subsets of $X$, then in $^*B = (^*X, ^*\mathcal{F}, ^*\mu)$, the set $^*\mathcal{F}$ is a collection of subsets of $^*X$; and if $\mu : \mathcal{F} \to \mathbb{R}$ then $^*\mu : {}^*\mathcal{F} \to {}^*\mathbb{R}$.

(iii) If $A$ is a set, from (ii) $^*A$ contains the set ${}^\sigma A = \{^*x : x \in A\}$, which is an elementwise copy of $A$. If $A$ is infinite, then ${}^*A \setminus {}^\sigma A \neq \varnothing$. It is possible to identify $A$ with ${}^\sigma A$ and then $^*A$ is an extension of $A$ just as $^*\mathbb{R}$ extends $\mathbb{R}$. (We will ensure that $^*r = r$ for $r \in \mathbb{R}$, so that $^*\mathbb{R} \supseteq \mathbb{R}$ always.) It is helpful to think of $^*A$ as an enrichment of $A$, containing new 'ideal' members, just as $^*\mathbb{R}$ is an enrichment of $\mathbb{R}$.

Other features of $^*V(\mathbb{R})$ will emerge in the detailed description below.

The nonstandard universe $^*V(\mathbb{R})$ can either be constructed or assumed as given and satisfying the axioms described below. The latter approach has the advantage that the important features are separated from details that are specific to a particular construction; moreover, the axiomatic approach eliminates the need to understand mathematical logic in order to utilise nonstandard methods. We now present the axioms for $^*V(\mathbb{R})$ as developed by Robinson and Zakon [**RZ**], Stroyan and Luxemburg [**SL**] and Keisler [**Ke$_2$**]. In the appendix we describe an algebraic construction of a nonstandard universe.

1.9. Axioms for $^*V(\mathbb{R})$.

*Axiom* I. There is a set $^*\mathbb{R} \supsetneq \mathbb{R}$; $^*\mathbb{R}$ is a set of atoms.[3]

*Axiom* II. There is a mapping $^* : V(\mathbb{R}) \to V(^*\mathbb{R})$ (sending $\mathbb{R}$ to $^*\mathbb{R}$) with the following properties:

(a) $^*r = r$ for $r \in \mathbb{R}$,

(b) (*Transfer principle*). For every $A_1, \ldots, A_n \in V(\mathbb{R})$ and every *bounded quantifier statement* $\phi(x_1, x_2, \ldots, x_n)$

$$\phi(A_1, \ldots, A_n) \text{ holds in } V(\mathbb{R}) \text{ if and only if } \phi(^*A_1, \ldots, ^*A_n) \text{ holds in } V(^*\mathbb{R}).$$

A *bounded quantifier statement* is one that can be made using only $\in$, $=$, &, or, not, $\Rightarrow$, and bounded quantifiers such as $\forall x \in y\ (\ldots)$ and $\exists x \in y\ (\ldots)$.

For the next axioms an important definition is needed: an object $A \in V(^*\mathbb{R})$ is said to be *internal* if $A \in {}^*B$ for some $B \in V(\mathbb{R})$; other objects in $V(^*\mathbb{R})$ are *external.*

*Axiom* III ($\omega_1$-*saturation*). If $A_1 \supseteq A_2 \supseteq \ldots$ are non-empty internal sets in $V(^*\mathbb{R})$ then $\bigcap_{n \in \mathbb{N}} A_n \neq \varnothing$.

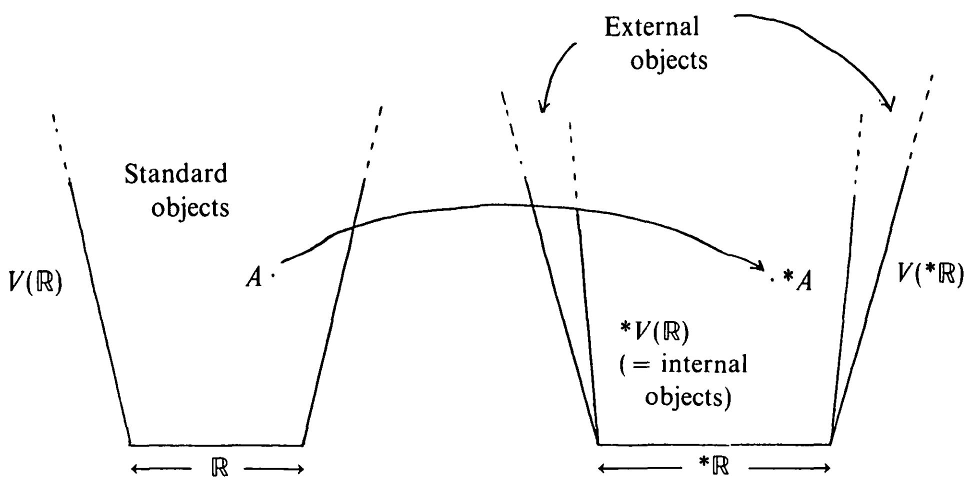
[3] $^*\mathbb{R}$ may be thought of as the hyperreals $^*\mathbb{R}$ as described (or constructed) earlier; but we are not making any such assumption a part of these axioms. Properties of $^*\mathbb{R}$ sketched earlier all follow from the later axioms.

*Axiom* IV (*Enlargement*). If $(A_j)_{j \in J}$ is a collecion of sets from $V(\mathbb{R})$ having the finite intersection property (f.i.p.) then $\bigcap_{j \in J} {}^*A_j \neq \varnothing$.

Given ${}^*\mathbb{R}$ and ${}^* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ satisfying the above axioms, ${}^*V(\mathbb{R})$ is defined by

$$\begin{aligned} {}^*V(\mathbb{R}) &= \{A : A \in {}^*B \text{ for some } B \in V(\mathbb{R})\} \\ &= \text{the collection of all } \textit{internal} \text{ objects in } V({}^*\mathbb{R}). \end{aligned}$$

The diagram below is a useful way to envisage $V(\mathbb{R})$, $V({}^*\mathbb{R})$ and ${}^*V(\mathbb{R})$.




Some discussion of these axioms is now in order.

Note first that quantification in mathematics is usually bounded (often only implicitly—we often write 'for all $x$' meaning 'for all $x \in \mathbb{R}$'), so most mathematical statements can be formalised as bounded quantifier statements (b.q.s's.) In practising nonstandard mathematics it is not necessary to actually formalise every statement in sight—only to develop the art of recognising an informal b.q.s. and the meaning of the transfer principle for it.

It is easy to show the following (noting first that $V_n(\mathbb{R}) \in V(\mathbb{R})$ and that $V_n(\mathbb{R})$ is transitive):

(i) ${}^*V_n(\mathbb{R})$ is transitive (that is, $C \in B \in {}^*V_n(\mathbb{R})$ implies $C \in {}^*V_n(\mathbb{R})$);

(ii) ${}^*V(\mathbb{R}) = \bigcup_{n \in \mathbb{N}} {}^*V_n(\mathbb{R})$; hence

(iii) if $A \in V(\mathbb{R})$ then ${}^*A$ is internal;

(iv) if $C \in B$ and $B$ is internal, then $C$ is internal.

Thus a b.q.s. $\phi({}^*A_1, \dots, {}^*A_n)$ talks only about internal objects, so these are the objects we know about via transfer; herein lies their importance.

Explicit constructs involving internal objects yield internal objects—via the transfer principle. For example, if $A$ and $B$ are internal then so are $A \cup B$, $A \cap B$, $A \times B$ etc. This idea is made precise and general as follows.

1.10. INTERNAL DEFINITION PRINCIPLE. *Let* $\phi(x, x_1, \ldots, x_n)$ *be a* b.q.s. *and let* $A, A_1, A_2, \ldots, A_n$ *be internal; then the set*

$$\{a \in A : \phi(a, A_1, \ldots, A_n) \textit{ holds in } V(^*\mathbb{R})\}$$

*is internal.*

*Proof.* Take $k$ with $A, A_1, \ldots, A_n \in {}^*V_k(\mathbb{R})$ and transfer the true b.q.s.:

'for all $B, B_1, \ldots, B_n \in V_k(\mathbb{R})$ there is $X \in V_k(\mathbb{R})$ such that $X \subseteq B$ and for all $b \in B$

$$b \in X \Leftrightarrow \phi(b, B)\text{'}.$$

The power set operation $\mathscr{P}$ has its counterpart $^*\mathscr{P}$ operating on internal sets. (Strictly we have $^*(\mathscr{P} \upharpoonright V_n(\mathbb{R}))$ for each $n$, and then we define $^*\mathscr{P}$ as the common extension of these internal functions.) If $A$ is internal and $A \in {}^*V_n(\mathbb{R})$ then by transfer

$$X \in {}^*\mathscr{P}(A) \Leftrightarrow X \subseteq A \textit{ and } X \in {}^*V_n(\mathbb{R}),$$

so $^*\mathscr{P}(A)$ is the set of all *internal* subsets of $A$. This means that in general $^*\mathscr{P}(A) \subsetneqq \mathscr{P}(A)$; so when interpreting the transfer of a statement about "all subsets" of a set we have to bear in mind that $^*\mathscr{P}(A)$ is meant, not $\mathscr{P}(A)$. This is one of the trickier points in nonstandard mathematics, and is where mistakes are easily made by the beginner. (On the other hand this is really what makes nonstandard methods work.) This point is illustrated by consideration of the 'completeness' of $^*\mathbb{R}$. The transfer of the completeness property for $\mathbb{R}$ tells us that any *internal* subset of $^*\mathbb{R}$ that is bounded above in $^*\mathbb{R}$ has a least upper bound. The set $\mathbb{R}$ is certainly bounded in $^*\mathbb{R}$, yet has no least upper bound; so we conclude that $\mathbb{R}$ itself is an *external* subset of $^*\mathbb{R}$. Similarly $\mathbb{N}$ and monad $(r)$ $(r \in \mathbb{R})$ are external sets.

The set of all functions from a set $X$ to a set $Y$, usually denoted $Y^X$, is a subset of $\mathscr{P}(X \times Y)$. From the above discussion we see that the internal counterpart of this operation on internal sets $A, B$ yields the set $\{f : f \text{ is } \textit{internal} \text{ and } f : A \to B\}$ which we will denote $B^A$ at the slight risk of confusion.

We can usefully formulate the axiom of $\omega_1$-saturation in the following alternative ways, where in each case $A$ is internal and $(A_n)_{n \in \mathbb{N}}$ is a sequence of internal sets.

1.11. PROPOSITION. *Axiom* III *is equivalent to each of* (a), (b), (c) *below.*

(a) *If* $\bigcap_{n \in \mathbb{N}} A_n \subseteq A$, *then* $\bigcap_{n \leqslant m} A_n \subseteq A$ *for some* $m \in \mathbb{N}$;

(b) *If* $A \subseteq \bigcup_{n \in \mathbb{N}} A_n$, *then* $A \subseteq \bigcup_{n \leqslant m} A_n$ *for some* $m \in \mathbb{N}$;

(c) *If* $(A_n)_{n \in \mathbb{N}}$ *is decreasing and* $\bigcap_{n \in \mathbb{N}} A_n = \varnothing$, *then* $A_m = \varnothing$ *for some* $m \in \mathbb{N}$.

The following useful axiom is often adopted instead of the axiom of $\omega_1$-saturation; the two axioms are in fact equivalent.

*Axiom* III′ (*Denumerable comprehension*). For every internal set $A$ and every function $f : \mathbb{N} \to A$ there is an *internal* function $g : {}^*\mathbb{N} \to A$ extending $f$.

Sometimes nonstandard methods require a saturation principle stronger than $\omega_1$-saturation. If $\kappa$ is a cardinal then the axiom of *$\kappa$-saturation* is given by:

*Axiom* $\mathrm{III}_\kappa$ (*$\kappa$-saturation*). If $(A_j)_{j\in J}$ is a collection of internal sets having f.i.p. and $|J| < \kappa$, then $\bigcap_{j\in J} A_j \neq \varnothing$.

In nonstandard measure theory it is mainly $\omega_1$-saturation that makes everything work; occasionally more saturation is required when topological considerations are involved.

C. Thompson has pointed out the following interesting consequence of $\omega_1$-saturation.

1.12. THEOREM. *If $(\delta_n)_{n\in\mathbb{N}}$ is a sequence from ${}^*\mathbb{R}^+$ with $\delta_n \to 0$ in the sense of ${}^*\mathbb{R}$, then $\delta_n$ is eventually $0$. Hence ${}^*\mathbb{R}$ is Cauchy-complete (and every ${}^*\mathbb{R}$-Cauchy sequence is eventually constant).*

*Proof.* We mean that given any $\varepsilon > 0$, $\varepsilon \in {}^*\mathbb{R}$, eventually $\delta_n < \varepsilon$. If $\delta_n$ is not eventually 0, choose a decreasing subsequence $\delta_{n_i}$ with $\delta_{n_i} > 0$, all $i$. Then $\bigcap_{i\in\mathbb{N}} (0, \delta_{n_i}) = \varnothing$, so by $\omega_1$-saturation, for some $k$ we have $(0, \delta_{n_k}) = \varnothing$, a contradiction.

Now take $\delta_n = |x_{n+1} - x_n|$ where $(x_n)_{n\in\mathbb{N}}$ is Cauchy in the sense of ${}^*\mathbb{R}$.

### §1(3) *Some basic nonstandard tools*

The following is a grab-bag of basic tools and ideas that will be referred to later.

(i) *Hyperfinite sets.* These are the internal sets corresponding to finite sets in $V(\mathbb{R})$.

1.13. DEFINITION. An internal set $A$ is *hyperfinite* (or **-finite*) if there is an internal bijection $f : \{1, 2, \ldots, H\} \to A$ for some (necessarily unique) $H \in {}^*\mathbb{N}$; we use $|A|$ to denote the internal cardinality $H$ of $A$.

The importance of infinite hyperfinite sets lies in the fact that they behave formally like finite sets (via the transfer principle); thus combinatorial and counting arguments valid for finite sets are valid for hyperfinite sets. For example, if $A$ is hyperfinite and $|A| = H$ then ${}^*\mathscr{P}(A)$ is hyperfinite and $|{}^*\mathscr{P}(A)| = 2^H$. A hyperfinite subset of an internal ordered set always has a maximum; a hyperfinite collection of hyperfinite sets has hyperfinite union; a hyperfinite set $X \subseteq {}^*\mathbb{R}$ can be summed, with sum written ${}^*\sum_{x\in X} x$. What this means is ${}^*\sum(X)$ where $\sum$ is the standard function in $V(\mathbb{R})$ that sums finite subsets of $\mathbb{R}$. Since ${}^*\sum$ extends $\sum$ it is usual to drop * from ${}^*\sum$.

Hyperfinite sets are useful for representing or carrying measures, since it is possible to perform integration by hyperfinite summation. Hyperfinite sets are also useful, as we shall see, for modelling large finite phenomena.

The next theorem is useful for finding hyperfinite representations of sets.

1.14. THEOREM. *Let $A$ be a standard set $(A \in V(\mathbb{R}))$; then there is a hyperfinite set $F$ with $A \subseteq F \subseteq {}^*A$ (where we are identifying $A$ with ${}^\sigma A$ so that $A \subseteq {}^*A$).*

*Proof.* Use the enlargement axiom on the sets $(D_x)_{x \in A}$ where $D_x = \{G : G \text{ finite}, x \in G \subseteq A\}$; taking $F \in \bigcap_{x \in A} {}^*D_x$ we have that $F$ is hyperfinite, $F \subseteq {}^*A$, and $x \in F$, all $x \in A$.

Note that this theorem does *not* tell us that every standard set is hyperfinite. In fact, unless $A$ is actually finite then $A$ (strictly, ${}^\sigma A$) is not internal, whereas hyperfinite sets are necessarily internal. (To see that ${}^\sigma A$ is not internal if $A$ is infinite, take an injection $f : \mathbb{N} \to A$. Then ${}^*f^{-1}({}^\sigma A) = \mathbb{N}$, which would thus be internal if $\sigma A$ were internal.)

(ii) *Overflow and underflow.* The *overflow principle* (and its contrapositive, *underflow*) is extremely useful; it is a simple consequence of the least number principle transferred to ${}^*\mathbb{N}$.

1.15. THEOREM. *Let $A$ be internal, $A \subseteq {}^*\mathbb{N}$.*

(a) Overflow principle. *If $n \in A$ for all finite $n \geqslant n_0$, then there is an infinite $H \in {}^*\mathbb{N}$ with $n \in A$ for all $n_0 \leqslant n \leqslant H$.*

(b) Underflow principle. *If $K \in A$ for all infinite $K \leqslant H$, for some infinite $H$, then there is finite $n_0$ such that $n \in A$ for all finite $n \geqslant n_0$.*

*Proof.* (a) If ${}^*\mathbb{N} \setminus A$ has infinite members let

$$H+1 = \min\{k : n_0 < k \in {}^*\mathbb{N} \setminus A\}$$

which exists because the set in question is internal.

(b) is the contrapositive of (a).

An important application of overflow is the following.

1.16. THEOREM. (a) (Robinson's sequential lemma). *If $(s_n)_{n \in {}^*\mathbb{N}}$ is internal and $s_n \approx 0$ for all finite $n$, there is infinite $H$ such that $s_n \approx 0$ for all $n \leqslant H$;*

(b) *Every countable sequence $(x_n)_{n \in \mathbb{N}}$ of infinitesimals has an infinitesimal bound;*

(c) *Every countable sequence $(y_n)_{n \in \mathbb{N}}$ of positive infinite hyperreals has an infinite lower bound.*

*Proof.* (a) Apply overflow to the set $A = \{n : |s_n| \leqslant 1/n\}$.

(b) Extend $(x_n)$ to $(x_n)_{n \in {}^*\mathbb{N}}$ using $\omega_1$-saturation and, taking $H$ as given by (a), let $b = \max\{|x_n| : n \leqslant H\}$.

(c) Apply (b) to $y_n^{-1}$.

Next we have an important application of underflow.

1.17. THEOREM (Infinitesimal overflow). *Suppose $S$ is internal, $S \subseteq {}^*\mathbb{R}$ and $x \in S$ whenever $0 < x \approx 0$; then for some real $\delta > 0$, ${}^*(0, \delta) \subseteq S$.*

*Proof.* Apply underflow to the set $A = \{K : (0, 1/K) \subseteq S\}$.

(iii) $\mathcal{S}$*-continuity.* Suppose that $g : {}^*[0, 1] \to {}^*\mathbb{R}$ is an internal function. There are two important notions of continuity applicable to $g$; the first is *continuity, which is the transfer of continuity. Thus $g$ is *continuous at $x \in {}^*[0, 1]$ if for any $\varepsilon > 0$, $\varepsilon \in {}^*\mathbb{R}$, there is $\delta > 0$, $\delta \in {}^*\mathbb{R}$ with $|g(y)-g(x)| < \varepsilon$ whenever $|y-x| < \delta$. The *continuous functions are precisely those in ${}^*C[0, 1]$, by transfer; for example $g(x) = \sin(xH)$ is *continuous ($H \in {}^*\mathbb{N}$).

The second notion, $\mathcal{S}$-continuity, allows us to relate internal functions back to the standard universe.

1.18. DEFINITION. Let $g : {}^*[0, 1] \to {}^*\mathbb{R}$; $g$ is $\mathcal{S}$*-continuous* if for all $x, y \in {}^*[0, 1]$, $g(x) \approx g(y)$ whenever $x \approx y$.

Note that if $g = {}^*f$, where $f \in C[0, 1]$, then ${}^*f$ is $\mathcal{S}$-continuous (Theorem 1.6); but there are $\mathcal{S}$-continuous functions not of this form, and not even in ${}^*C[0, 1]$. For example, putting $\Delta t = H^{-1}$ for infinite $H$, let

$$g(x) = \begin{cases} \Delta t & \text{if } 2k\Delta t \leqslant x < (2k+1)\Delta t \\ -\Delta t & \text{if } (2k+1)\Delta t \leqslant x < (2k+2)\Delta t \end{cases} \qquad (k \in {}^*\mathbb{N}).$$

Then $g$ is $\mathcal{S}$-continuous, but not *continuous. Notice that the function $\sin(xH)$ is *continuous but not $\mathcal{S}$-continuous if $H$ is infinite, so there is no implication between these two notions.

The important result that relates $\mathcal{S}$-continuity to standard continuity is the following.

1.19. THEOREM. *Let $g : {}^*[0, 1] \to {}^*\mathbb{R}$ be internal and $\mathcal{S}$-continuous, with $g(0)$ finite. Then there is a unique function $f \in C[0, 1]$ such that $f(x) = {}^\circ g(x)$ for all $x \in [0, 1]$.*

*Proof.* Define $f : [0, 1] \to \overline{\mathbb{R}}$ by $f(x) = {}^\circ g(x) = {}^\circ g(y)$ for any $y \approx x$. Take real $\varepsilon > 0$. Applying Theorem 1.17 to the set $\{\alpha : |x-y| < \alpha \Rightarrow |g(x)-g(y)| < \varepsilon\}$, we find real $\delta > 0$ such that $|g(x)-g(y)| < \varepsilon$ whenever $|x-y| < \delta$. So $f(x)$ is finite for all $x$, and $f$ is continuous on $[0, 1]$.

The standard function $f$ obtained from the $\mathcal{S}$-continuous internal function $g$ is written $f = {}^\circ g$.

The notion of $\mathcal{S}$-continuity extends in the obvious way to functions defined on other subsets of ${}^*\mathbb{R}$. An important instance of this is a function $g : T \to {}^*\mathbb{R}$ where $T$ is the discrete time set $\{0, \Delta t, 2\Delta t, \ldots, 1\}$, with $\Delta t = H^{-1}$ for some infinite $H$. Such a function $g$ is said to be $\mathcal{S}$-continuous if $g(s) \approx g(t)$ whenever $s \approx t$, $s, t \in T$; and if $g$ is $\mathcal{S}$-continuous and $g(0)$ is finite there is a unique $f \in C[0, 1]$ given by $f(x) = {}^\circ g(t)$ for any $t \approx x$, $t \in T$. This can be established immediately from Theorem 1.19 by filling $g$ in linearly to make an $\mathcal{S}$-continuous function $\hat{g} : {}^*[0, 1] \to {}^*\mathbb{R}$.

(iv) *Convention.* We mentioned earlier that it is possible to identify a set $A$ with the subset ${}^\sigma A = \{{}^*x : x \in A\}$ of ${}^*A$, so that $A$ may be regarded as embedded in the enrichment ${}^*A$ of $A$. This will be done frequently and implicitly in the sequel when discussing a particular set (which may be the carrier for a topology or measure etc.)

Suppose that $B$ is another set and we regard $B \subseteq {}^*B$ also; consider a function $f : A \to B$ and the corresponding ${}^*f : {}^*A \to {}^*B$. Then ${}^*f$ is an extension of $f$, since for $x \in A$, $y \in B$, if $f(x) = y$, then ${}^*f({}^*x) = {}^*y$ by transfer, so ${}^*f(x) = y$ under the identification of $x$, $y$ with ${}^*x$ and ${}^*y$. In such a situation it is sometimes convenient to drop * from ${}^*f$, at the risk of some ambiguity. If this is done the context will make things clear.

## §1(4) *Nonstandard topology*

We shall need a few basics from nonstandard topology, beginning with the generalisation of the idea of the monad of an element. For proofs and further information consult any of the basic books [**D**, **Rb**, **SL**].

1.20. DEFINITIONS. (i) Let $(X, \tau)$ be a topological space, and let $x \in X$. The *monad* of $x$ is defined by

$$\operatorname{monad}(x) = \bigcap_{x \in V \in \tau} {}^*V .$$

(ii) Let $z \in {}^*X$ and let $x \in X$; write $z \approx x$ if $z \in \operatorname{monad}(x)$; then $z$ is said to be *nearstandard* if $z \approx x$ for some $x \in X$. The set of nearstandard points in ${}^*X$ is denoted $\operatorname{ns}({}^*X)$. (Note that $\approx$ is not a symmetric relation in this context.)

The Hausdorff property is easily characterised using monads.

1.21. THEOREM. *The space $(X, \tau)$ is Hausdorff if and only if for all $x \neq y \in X$, $\operatorname{monad}(x) \cap \operatorname{monad}(y) = \varnothing$. In this case, if $z \in \operatorname{ns}({}^*X)$, there is a unique $x \in X$ with $z \approx x$.*

1.22. DEFINITION. If $(X, \tau)$ is Hausdorff and $z \in \operatorname{ns}({}^*X)$, the *standard part* of $z$ (written ${}^\circ z$ or $\operatorname{st}(z)$) is the unique $x \in X$ with $z \approx x$.

The nonstandard characterisation of continuity generalises in the obvious way.

1.23. THEOREM. *Let $(X, \tau)$ and $(X', \tau')$ be topological spaces, and let $f : X \to X'$. Then $f$ is continuous at $x \in X$ if and only if $f(z) \approx f(x)$ whenever $z \approx x$; that is, $f(\operatorname{monad}(x)) \subseteq \operatorname{monad}(f(x))$.*

(*N.B.* We are here adopting the convention discussed in §1(3)(iii) above, so that $X \subseteq {}^*X$, $X' \subseteq {}^*X'$ and $f$ is really ${}^*f$ in the last line.)

We continue with a brief discussion of compactness.

1.24. THEOREM. *Let $(X, \tau)$ be a topological space; then $X$ is compact if and only if every point in ${}^*X$ is nearstandard.*

The following helps to explain why finite approximation arguments work for compact spaces.

1.25. Corollary. *If* $(X, \tau)$ *is compact, there is a hyperfinite set* $F$ *with* $X \subseteq F \subseteq \mathrm{ns}({}^*X)$, *such that* $X = \mathrm{st}(F)$.

*Proof.* By Theorem 1.24, $\mathrm{ns}({}^*X) = {}^*X$; apply Theorem 1.14.

With $\kappa$-saturation (where $\tau$ has a base of cardinality $< \kappa$) the following useful results obtain.

1.26. Theorem. *Let* $(X, \tau)$ *be a Hausdorff topological space.*

(a) *If* $A$ *is internal,* $A \subseteq {}^*X$, *then* $\mathrm{st}(A)$ *is closed in* $X$.

(b) *If* $X$ *is regular, and* $A$ *is internal with* $A \subseteq \mathrm{ns}({}^*X)$, *then* $\mathrm{st}(A)$ *is compact.*

(Note that (b) gives the converse to Corollary 1.25 for regular spaces.)

Consider now a product space $(X, \tau) = \left(\prod_{i \in I} X_i, \prod_{i \in I} \tau_i\right)$ where $(X_i, \tau_i)$ are topological spaces. We have ${}^*X = \prod_{i \in {}^*I} {}^*X_i$ = the *internal* set of all *internal* functions $x : {}^*I \to \bigcup_{i \in {}^*I} {}^*X_i$ such that $x_i \in {}^*X_i$ for all $i$. Note however that nearstandardness depends only on $x_i$ for $i \in I$.

1.27. Proposition. *For a product space* $X$ *as above,* $z \in \mathrm{ns}({}^*X)$ *if and only if* $z_i \in \mathrm{ns}({}^*X_i)$ *for all* $i \in I$.

This fact together with Theorem 1.24 gives a simple proof of Tychonoff's theorem.

## §1(5) *Metric spaces*

For a nonstandard treatment of a metric space $(M, d)$ we specialise the preceding theory, or we can think of directly generalising from the metric space $(\mathbb{R}, |\cdot|)$. On ${}^*M$ we have the metric function ${}^*d : {}^*M \times {}^*M \to {}^*\mathbb{R}^+$ satisfying the metric axioms. We usually omit $*$ from ${}^*d$. For $x, y \in {}^*M$ we define the symmetric relation $\approx$ by $x \approx y \Leftrightarrow d(x, y) \approx 0$. For $x \in M$, $\mathrm{monad}(x) = \{z \in {}^*M : z \approx x\}$ and for $z \in {}^*M$, $z \approx x$ has the same meaning as the topological definition earlier. The set ns(${}^*M$) and the standard part function $^\circ = \mathrm{st} : \mathrm{ns}({}^*M) \to M$ are defined as before. For a metric space Theorem 1.26 needs only $\omega_1$-saturation.

## §1(6) *Measure spaces*

Since the rest of this survey has to do with nonstandard measure theory, we restrict ourselves here to a few opening remarks.

Suppose that $\mathbf{X} = (X, \mathscr{F}, \mu)$ is a standard measure space with $\mu$ finite; then ${}^*\mathbf{X} = ({}^*X, {}^*\mathscr{F}, {}^*\mu)$ is an internal measure space. This means that ${}^*\mathscr{F} \subseteq {}^*\mathscr{P}({}^*X)\ (\subseteq \mathscr{P}({}^*X))$ and ${}^*\mu : {}^*\mathscr{F} \to {}^*\mathbb{R}^+$. Now $\mathscr{F}$ is a $\sigma$-algebra; the transfer of this means that ${}^*\mathscr{F}$ is a ${}^*\sigma$-algebra (*not* that ${}^*\mathscr{F}$ is a $\sigma$-algebra); thus, if $(A_n)_{n \in {}^*\mathbb{N}}$ is an internal sequence of internal sets from ${}^*\mathscr{F}$, then $\bigcup_{n \in {}^*\mathbb{N}} A_n \in {}^*\mathscr{F}$.

Could ${}^*\mathscr{F}$ nevertheless be a $\sigma$-algebra? Using $\omega_1$-saturation (1.11(b)) we see that

$\bigcup_{n \in \mathbb{N}} A_n \notin {}^*\mathscr{F}$ unless there is $m$ with $\bigcup_{n \in \mathbb{N}} A_n = \bigcup_{n \leq m} A_n$. So unless $\mathscr{F}$ is finite, ${}^*\mathscr{F}$ is *not* a $\sigma$-algebra; this seemed for a long time to be the sticking point for a good nonstandard approach to measure theory. Curiously, however, it is precisely this point that allows the successful Loeb measure spaces to be defined—as we shall see in §3.

Turning to ${}^*\mu$, we see that ${}^*\mu$ is ${}^*\sigma$-additive (*not* $\sigma$-additive) on ${}^*\mathscr{F}$; that is, if $(A_n)_{n \in {}^*\mathbb{N}}$ is an internal sequence of disjoint sets from ${}^*\mathscr{F}$, then ${}^*\mu\left(\bigcup_{n \in {}^*\mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} {}^*\mu(A_n)$, where the meaning of the infinite sum as a limit of a ${}^*\mathbb{N}$-indexed sequence of hyperreals is the transfer of the corresponding standard definition.

Associated with ${}^*\mathbf{X}$ we have the ${}^*\mathscr{F}$-measurable functions (which are those internal functions ${}^*X \to {}^*\overline{\mathbb{R}}$ satisfying the transfer of the definition of measurability), and the subset ${}^*(\mathscr{L}^1(\mathbf{X}))$ of *integrable functions. For each $g \in {}^*(\mathscr{L}^1(\mathbf{X}))$ we have ${}^*\int g\, d{}^*\mu$, by which is meant the operation ${}^*\int$ applied to $g$ and ${}^*\mu$.

Much of the power of Loeb measures stems from the fact that there are many internal measure spaces that do not derive from a standard space as does $({}^*X, {}^*\mathscr{F}, {}^*\mu)$ above. Saying that $\mathbf{Y} = (Y, \mathscr{G}, \nu)$ is an internal measure space (really a *measure space) means that $\mathbf{Y}$ is internal and satisfies the transfer of the requirements for a measure space, that is, $\mathscr{G} \subseteq {}^*\mathscr{P}(Y)$, $\mathscr{G}$ is a ${}^*\sigma$-algebra and $\nu : \mathscr{G} \to {}^*\mathbb{R}^+$ is ${}^*\sigma$-additive.

For an internal measure space $\mathbf{Y}$ we have the notion of $\mathscr{G}$-measurable functions and the subset ${}^*\mathscr{L}^1(\mathbf{Y})$ of these. What is meant here is that ${}^*\mathscr{L}^1(\cdot)$ is the *image of the operation $\mathscr{L}^1(\cdot)$ that applies to standard measure spaces; equivalently, ${}^*\mathscr{L}^1(\mathbf{Y})$ is the set of those functions $Y \to {}^*\overline{\mathbb{R}}$ that satisfy the transfer of the standard definition of integrability. We have for $g \in {}^*\mathscr{L}^1(\mathbf{Y})$ the integral ${}^*\int g\, d\nu \in {}^*\mathbb{R}$, which is the operator ${}^*\int$ applied to $g$ and $\nu$. It is customary to drop the * from ${}^*\int$.

Note that if $\mathbf{Y} = {}^*\mathbf{X}$, as considered earlier, then we have the sets ${}^*(\mathscr{L}^1(\mathbf{X}))$ and ${}^*\mathscr{L}^1({}^*\mathbf{X})$ (that is, ${}^*\mathscr{L}^1(\cdot)$ applied to ${}^*\mathbf{X}$), which are the same, by transfer; we shall write ${}^*\mathscr{L}^1(\mathbf{X})$ for this set.

## §2 Early Nonstandard Measure Theory

'Early' efforts to introduce nonstandard methods into measure theory met with only modest success. In the main they gave methods of representing measures already constructed, and provided pleasant intuitive nonstandard counterparts to the standard theory. The obstacle to developing powerful applications of nonstandard methods in measure theory was that measures constructed were (with a single exception) only finitely additive, because infinite nonstandard fields of sets refuse to be $\sigma$-additive (although they may be ${}^*\sigma$-additive). The way in which this obstacle was surmounted—indeed exploited—with the introduction of Loeb measure is described in §3.

In this section we give a brief sketch of results and methods developed prior to Loeb measure; proofs are omitted.

### §2(1) *Lebesgue measure*

Robinson's treatment **[Rb]** of Lebesgue measure $m$ on $\mathbb{R}$ is the natural nonstandard parallel to the standard Lebesgue theory.

2.1. THEOREM (Robinson). *Let* $A \subseteq \mathbb{R}$; *then the following are equivalent.*

(a) *$A$ is Lebesgue measurable;*

(b) *there are sets* $F, G \subseteq {}^*\mathbb{R}$, **closed and *open respectively, such that* $F \subseteq {}^*A \subseteq G$ *and* ${}^*m(F) \approx {}^*m(G)$, *with* ${}^*m(F)$ *finite.*

If (b) *holds then* $m(A) = {}^\circ({}^*m(F)) = {}^\circ({}^*m(G))$.

This leads to a natural representation of the Lebesgue integral as follows, restricting to $[0, 1]$ for convenience.

2.2. THEOREM. *Let* $f : [0, 1] \to \mathbb{R}$ *be Lebesgue measurable with* $|f(x)| \leqslant K$ *for all* $x$. *Let* $-K = y_0 < y_1 < \ldots < y_H = K$, *be a hyperfinite partition of* ${}^*[-K, K]$, *with* $y_{i+1} - y_i \approx 0$ *for all* $i$. *Then*

$$\int f dm = {}^\circ\sum_{i=1}^{H} y_i {}^*m(\{x : y_{i-1} \leqslant {}^*f(x) < y_i\}).$$

Robinson suggested that these results can be used to provide the foundation of Lebesgue theory. For measure, it is necessary to assume that $m$ has already been defined for open and closed sets, and then use Theorem 2.1. The integral (for bounded functions) may be *defined* by the equation in Theorem 2.2 after showing that the sum on the right is independent of the partition of $[-K, K]$.

A more intuitive representation of Lebesgue measure was developed by Bernstein and Wattenberg [**BW**] using the following counting approach.

2.3. DEFINITION. Let $X$ be any set and let $S$ be a hyperfinite subset of ${}^*X$; $S$ is called a *sample*. For internal $A \subseteq {}^*X$ define

$$\nu_S(A) = \frac{|A \cap S|}{|S|}$$

(where $|\cdot|$ denotes internal cardinality). Clearly $\nu_S$ is an internal totally additive probability measure on the internal subsets of ${}^*X$; if we define ${}^\circ\nu_S$ by ${}^\circ\nu_S(A) = {}^\circ(\nu_S(A))$ then ${}^\circ\nu_S$ is a standard finitely additive probability measure[(4)] on the internal subsets of ${}^*X$.

Lebesgue measure on $[0, 1]$ can be obtained if $S$ is chosen carefully.

2.4. THEOREM [**BW**]. *There is a sample* $S \subseteq {}^*[0, 1]$ *such that for every Lebesgue measurable set* $B \subseteq [0, 1]$

$$m(B) = {}^\circ\nu_S({}^*B) = {}^\circ(|{}^*B \cap S|/|S|).$$

A sample $S \subseteq {}^*[0, 1]$ satisfying this theorem is called a *Lebesgue sample.* Note that *every* subset $B$ of $[0, 1]$ is assigned a measure by ${}^\circ\nu_S({}^*B)$; so ${}^\circ\nu_S$ gives a finitely additive extension of $m$ to all subsets of $[0, 1]$. The sample obtained in [**BW**] has the following additional features.

---

(4) What we mean is that setting $\mathbf{Y} = (Y, \mathscr{A}, \mu)$ where $Y = {}^*X$, $\mathscr{A}$ = the internal subsets of ${}^*X$ and $\mu = {}^\circ\nu_S$ then $\mathscr{A}$ is an algebra on $Y$, $\mu : \mathscr{A} \to [0, 1]$ and $\mu$ is finitely additive; so $\mathbf{Y}$ is a finitely additive probability space in the standard sense. We are not claiming that $\mathbf{Y} \in V(\mathbb{R})$; $\mathbf{Y}$ may lie outside $V(\mathbb{R})$.

(i) $\mathbb{R} \subseteq S$; this means that if $A \neq \emptyset$ then $\nu_S(^*A) \neq 0$, though of course if $A$ is null then $\nu_S(^*A) \approx 0$.

(ii) $S+q = S(\bmod 1)$ for all rationals $q$. Thus, for example, $\nu_S(^*([0,\frac{1}{4}) \cap \mathbb{Q})) = \frac{1}{2}\nu_S(^*([0,\frac{1}{2}) \cap \mathbb{Q}))$, which is a way of making precise the intuition that "the probability of hitting a rational number in the interval $[0,\frac{1}{4})$ is exactly half that of hitting a rational number in the interval $[0,\frac{1}{2})$, despite the fact that both sets in question have Lebesgue measure zero" [**BW**].

(iii) For any real number $r$, and internal $A \subseteq {}^*[0,1]$, $\nu_S(A+r) \approx \nu_S(A)$ (where $+$ is modulo 1). This means that $^\circ\nu_S(^*\cdot)$ is a *translation invariant* finitely additive extension of $m$.

It is worth noting some interesting alternative constructions of Lebesgue samples that were given in [$\mathbf{W_2}$].

2.5. DEFINITION. Let $(X, \mathcal{F}, \mu)$ be a standard measure space; an element $x \in {}^*X$ is *random* if $x \notin {}^*N$ whenever $N$ is $\mu$-null $(N \subseteq X)$.

By the enlargement axiom, there is a *null set $Y \subseteq {}^*X$ with $Y \supseteq {}^*N$ for all null $N$, so almost all points in $^*X$ are random.

2.6. THEOREM [$\mathbf{W_2}$]. (a) *Let $x \in {}^*([0,1]^{\mathbb{N}})$ be random with respect to the usual product probability measure on $[0,1]^{\mathbb{N}}$. Then there is an infinite $H_0 \in {}^*\mathbb{N}$ such that $S = \{x_1, x_2, \ldots, x_H\}$ is a Lebesgue sample whenever $H \geqslant H_0$.*

(b) *Let $x \in {}^*[0,1]$ be random (with respect to Lebesgue measure); take $\alpha$ irrational and infinite $H \in {}^*\mathbb{N}$. Then $S = \{x+k\alpha : k = 0,1,2,\ldots,H\}$ is a Lebesgue sample (addition is again modulo 1).*

(*N.B.* For (a) the strong law of large numbers is invoked, and for (b) the ergodic theorem.)

A Lebesgue sample $S$ gives a representation of the Lebesgue integral in the expected way.

2.7. THEOREM [**BW**]. *If $f : [0,1] \to \mathbb{R}$ is bounded and Lebesgue measurable, then*

$$\int f\,dm = {}^\circ\sum_{x \in S} {}^*f(x)\Delta\nu_S$$

*where $\Delta\nu_S = |S|^{-1}$. If $S$ is chosen with care* [**Be**], *this holds for all $f \in \mathcal{L}^1[0,1]$.*

Wattenberg [$\mathbf{W_1}$] extended the sample approach to give a similar and simultaneous representation of Hausdorff measures $\lambda^t$ (for $t \in [0,\infty)$) on a metric space.

## §2(2) *Finitely additive measures*

The sample representation was extended to arbitrary non-atomic finitely additive probability measures by Henson [$\mathbf{He_1}$], who also used this idea to give a simple general method of constructing invariant finitely additive measures. Hyperfinite samples were used by Parikh and Parnes [**PP**] to obtain finitely additive conditional probability measures on the reals.

In a series of papers [$\mathbf{Lo_1}$, $\mathbf{Lo_2}$, $\mathbf{Lo_3}$] Loeb developed a fresh approach to the nonstandard representation of finite finitely additive signed measures on a given measurable space $(X, \mathscr{F})$. From one point of view this generalised the sample approach by using a weighted sample.

With $(X, \mathscr{F})$ fixed ($X$ infinite, $\mathscr{F}$ a $\sigma$-field) use Theorem 1.14 to obtain a hyperfinite algebra $\mathscr{G}$ with $\mathscr{F} \subseteq \mathscr{G} \subseteq {}^*\mathscr{F}$, and thus a hyperfinite partition $\{A_1, \ldots, A_H\}$ of ${}^*X$ into disjoint subsets such that

(i) $A_i \in {}^*\mathscr{F}$ for all $i$;

(ii) if $B \in \mathscr{F}$ then ${}^*B = \bigcup_{A_i \subseteq {}^*B} A_i$.

We illustrate what can now be obtained.

2.8. THEOREM. *Let $\mu$ be a finite, finitely additive signed measure on $(X, \mathscr{F})$.*

(a) *If $B \in \mathscr{F}$, then $\mu(B) = \sum_{A_i \subseteq {}^*B} {}^*\mu(A_i)$ and $|\mu|(B) = \sum_{A_i \subseteq {}^*B} |{}^*\mu(A_i)|$.*

(b) *If $X$ has a Hahn decomposition $X = B_+ \cup B_-$ with respect to $\mu$, then ${}^*B_+ \supseteq \bigcup_{{}^*\mu(A_i) > 0} A_i$ and ${}^*B_- \supseteq \bigcup_{{}^*\mu(A_i) < 0} A_i$.*

(c) *A bounded function $f : X \to \mathbb{R}$ is measurable if and only if for each $i$, ${}^*f(x) \approx {}^*f(y)$ whenever $x, y \in A_i$, and if so then $\int f d\mu = {}^\circ\sum {}^*f(a_i){}^*\mu(A_i)$ where $\{a_i\}_{i \leqslant H}$ is any internal sequence with $a_i \in A_i$ for each $i$.*

To see that this can be viewed as a weighted sample representation of $\mu$, write $S = \{a_1, \ldots, a_H\}$ and $\Delta{}^*\mu(a_i) = {}^*\mu(A_i)$; then (a), (c) above can be written

$$\mu(B) = \sum_{x \in {}^*B \cap S} \Delta{}^*\mu(x) \quad \text{and} \quad \int f d\mu = {}^\circ\sum_{x \in S} {}^*f(x)\Delta{}^*\mu(x).$$

It was shown in [**BL**] how to extend 2.8(c) above to unbounded functions by choosing the partition of ${}^*X$ carefully (for a particular measure $\mu$).

Loeb [$\mathbf{Lo_2}$] used the above framework to give among other things an intuitive proof of a representation theorem for the dual of $\mathscr{L}^\infty$, due to Yosida and Hewitt. Suppose that $\mathscr{N} \subseteq \mathscr{F}$ and $\mathscr{N}$ is closed under countable unions, and subsets from $\mathscr{F}$. (Think of $\mathscr{N}$ as a collection of null sets.) Let $\mathscr{L}^\infty$ be the set of measurable functions on $X$ that are bounded off some set in $\mathscr{N}$, and let $\mathscr{L}^{\infty(d)}$ denote the dual of $\mathscr{L}^\infty$. Finally, let $\mathscr{M}$ be the Banach space of finite finitely additive signed measures that are zero on sets in $\mathscr{N}$, with norm $\|\mu\| = |\mu|$.

Set $I_1 = \{i : A_i \notin {}^*\mathscr{N}\}$ and let $E = {}^*\mathbb{R}^{I_1}$; then $E$ is a hyperfinite dimensional ${}^*$Euclidean space (and it is also a vector space over $\mathbb{R}$). Each of $\mathscr{L}^\infty$, $\mathscr{L}^{\infty(d)}$ and $\mathscr{M}$ has a natural representation in $E$ as follows. Let $f \in \mathscr{L}^\infty$, $\phi \in \mathscr{L}^{\infty(d)}$ and $\mu \in \mathscr{M}$ and define $\bar{f}, \bar{\phi}, \bar{\mu} \in E$ by

$$\bar{f}_i = {}^*f(a_i),$$

$$\bar{\phi}_i = {}^*\phi(\chi_{A_i}), \qquad (\chi_{A_i} \text{ is the indicator function of } A_i),$$

$$\bar{\mu}_i = {}^*\mu(A_i).$$

It is quite straightforward to establish the following facts.

2.9. THEOREM.

(a) $\phi(f) \approx \sum_{i \in I_1} \bar{\phi}_i \bar{f}_i = (\bar{\phi}, \bar{f})$ *(the inner product in E)*;

(b) $\int f d\mu \approx \sum_{i \in I_1} \bar{\mu}_i \bar{f}_i = (\bar{\mu}, \bar{f})$;

(c) $\|f\|_\infty \approx \max_{i \in I_1} |\bar{f}_i|$;

(d) $\|\phi\| \approx \sum_{i \in I_1} |\bar{\phi}_i|$;

(e) $\|\mu\| \approx \sum_{i \in I_1} |\bar{\mu}_i|$.

It is now easy to establish the Yosida–Hewitt representation result, namely that $\mathscr{L}^{\infty(d)}$ and $\mathscr{M}$ are isometrically isomorphic as Banach spaces. For $\phi \in \mathscr{L}^{\infty(d)}$ define $\mu_\phi \in \mathscr{M}$ by $\mu_\phi(B) = \phi(\chi_B)$. Theorem 2.9 shows that the mapping $\phi \to \mu_\phi$ is an isometry with inverse given by $\mu \to \phi_\mu$ where $\phi_\mu(f) = \int f d\mu$.

In the same paper [Lo$_2$] Loeb gave a construction of the Stone space $\hat{X}$ (associated with $(X, \mathscr{F})$ and $\mathscr{N}$) as $I_1/\sim$ where $i \sim j$ if $^*f(a_i) \approx {}^*f(a_j)$ for all $f \in \mathscr{L}^\infty$. Then $\hat{X}$ embeds naturally in $\mathscr{L}^{\infty(d)}$ by $(i/\sim)(f) = {}^\circ(\bar{f}_i)$; it is easy to establish that $\hat{X}$ is compact (in the relative weak topology induced by $\mathscr{L}^\infty$) and that $\mathscr{L}^\infty \cong C(\hat{X})$. (Strictly this should be $L^\infty$ of course.)

## §3 LOEB MEASURES

Nonstandard measure theory received a fresh impetus with the discovery of the Loeb construction in [Lo$_4$], which is a very simple (with hindsight) application of $\omega_1$-saturation. The idea is to take an internal measure space (which will in general be only *finitely* additive, and need not derive from any standard space) and from it construct a new *standard*, $\sigma$-additive measure space. Loeb used to advantage the fact that infinite internal algebras are never $\sigma$-fields, which seemed to be the obstacle to a useful nonstandard measure theory. The resulting *Loeb spaces* are rich and varied, and their range of applicability will be discussed in later sections. Here we begin with the construction of Loeb measures, and the development of the Loeb theory; proofs of basic results are provided in this section.

### §3(1) *Construction of Loeb measures*

Recall from §1(6) that by an internal finitely additive measure space $\mathbf{Y} = (Y, \mathscr{A}, \nu)$ we mean the following:

(i) $Y, \mathscr{A}, \nu$ are all internal and $\mathscr{A} \subseteq {}^*\mathscr{P}(Y)$,

(ii) $\mathscr{A}$ is an algebra,

(iii) $\nu : \mathscr{A} \to {}^*\mathbb{R}^+$ and $\nu(A \cup B) = \nu(A) + \nu(B)$ if $A \cap B = \varnothing$.

3.1. THEOREM (Loeb measure [$\mathbf{Lo_4}$, $\mathbf{Lo_8}$]). *Let* $\mathbf{Y} = (Y, \mathscr{A}, \nu)$ *be an internal finitely additive measure space. Define* $°\nu : \mathscr{A} \to \bar{\mathbb{R}}^+$ *by* $°\nu(A) = °(\nu(A))$. *Then* $°\nu$ *has a unique* $\sigma$*-additive extension* $\lambda$ *to* $\sigma(\mathscr{A})$, *the* $\sigma$*-algebra generated by* $\mathscr{A}$, *which is given by* $\lambda(B) = \inf\{°\nu(A) : B \subseteq A, A \in \mathscr{A}\}$. *Further, if* $\lambda(B) < \infty$, *then* $\lambda(B) = \sup\{°\nu(A) : A \subseteq B, A \in \mathscr{A}\}$ *and there is* $A \in \mathscr{A}$ *with* $\lambda(A \,\Delta\, B) = 0$.

*Proof.* Note that $°\nu$ is obviously finitely additive; in [$\mathbf{Lo_4}$] the existence of $\lambda$ is obtained by appealing to the Carathéodory extension theorem [**Ry**] which requires the following to be verified: if $(A_n)_{n \in \mathbb{N}}$ is an increasing sequence from $\mathscr{A}$ with $\bigcup_{n \in \mathbb{N}} A_n = A \in \mathscr{A}$, then $°\nu(A_n) \to °\nu(A)$. Now $\omega_1$-saturation (1.11(b)) means that if $\bigcup_{n \in \mathbb{N}} A_n = A \in \mathscr{A}$ then $A = A_m$ for some $m$, so the condition is verified trivially.

It is actually very easy to see from first principles that $\lambda$ as defined in the theorem is $\sigma$-additive on $\sigma(\mathscr{A})$. First show that $\lambda$ is finitely additive. Then consider a countable sequence $B_n$ of mutually disjoint sets in $\sigma(\mathscr{A})$. Clearly $\sum \lambda(B_n) \leqslant \lambda(\bigcup B_n)$, and if $\sum \lambda(B_n) = \infty$ we are done. For the reverse inequality when $\sum \lambda(B_n) = s < \infty$, take $\varepsilon > 0$ and for $n \in \mathbb{N}$ choose $A_n \in \mathscr{A}$ with $B_n \subseteq A_n$ and $\nu(A_n) < \lambda(B_n) + \varepsilon 2^{-n}$. Extend $(A_n)_{n \in \mathbb{N}}$ to $(A_n)_{n \in {}^*\mathbb{N}}$ (using denumerable comprehension) so that $A_n \in \mathscr{A}$ for all $n \in {}^*\mathbb{N}$. For each finite $m$

$$\nu\left(\bigcup_{n \leqslant m} A_n\right) < \sum_{n \leqslant m} \left(\lambda(B_n) + \varepsilon 2^{-n}\right) < s + \varepsilon$$

so by overflow there is infinite $H$ with

$$\nu\left(\bigcup_{n \leqslant H} A_n\right) < s + \varepsilon .$$

Putting $A = \bigcup_{n \leqslant H} A_n \in \mathscr{A}$, we have $\bigcup_{n \in \mathbb{N}} B_n \subseteq A$ and $°\nu(A) \leqslant s + \varepsilon$, so $\lambda\left(\bigcup_{n \in \mathbb{N}} B_n\right) \leqslant s$, as required.

Uniqueness of $\lambda$ follows easily when $\lambda(Y) < \infty$, by taking complements. For the case when $\lambda(Y) = \infty$, uniqueness was established by Henson [$\mathbf{He_3}$] who showed that for $B \in \sigma(\mathscr{A})$ *either* there is $A \in \mathscr{A}$ with $A \subseteq B$ and $°\nu(A) = \infty$ *or* there is a sequence $(A_n)_{n \in \mathbb{N}}$ from $\mathscr{A}$ with $B \subseteq \bigcup A_n$ and $°\nu(A_n) < \infty$ for all $n$. In either case $\lambda(B)$ is uniquely determined.

Now suppose that $\lambda(B) < \infty$; then there is an internal set $D \supseteq B$ with $\nu(D)$ finite, so the inner approximation $\lambda(B) = \sup\{°\nu(A) : A \subseteq B, A \in \mathscr{A}\}$ is obtained by taking complements with respect to $D$. Now let $(A_n)_{n \in \mathbb{N}}$ be an increasing sequence of sets from $\mathscr{A}$ with $A_n \subseteq B$ and $|\nu(A_n) - \lambda(B)| < 1/n$; by denumerable comprehension and overflow there is infinite $H \in {}^*\mathbb{N}$ with $|\nu(A_H) - \lambda(B)| < 1/H$, and such that $A_H \in \mathscr{A}$ and $A_H \supseteq A_n$ for all $n \leqslant H$. Then $°\nu(A_H) = \lambda(B)$ and it is easy to see that $\lambda(A_H \,\Delta\, B) = 0$.

It has been customary to call the completion of the measure space $(Y, \sigma(\mathscr{A}), \lambda)$ the *Loeb space* of $\mathbf{Y}$, written $L(\mathbf{Y})$, as follows.

3.2. DEFINITION. Let $\mathbf{Y} = (Y, \mathscr{A}, \nu)$ be an internal finitely additive measure space. The *Loeb space* $L(\mathbf{Y})$ is the standard measure space $L(\mathbf{Y}) = (Y, L(\mathscr{A}), L(\nu))$ where

(i) $L(\mathscr{A})$ is the completion of $\sigma(\mathscr{A})$ with respect to the measure $\lambda$ given by Theorem 3.1; $L(\mathscr{A})$ is the *Loeb $\sigma$-algebra* (associated with $\mathscr{A}$ and $\nu$); sets in $L(\mathscr{A})$ are called *Loeb measurable*;

(ii) $L(\nu)$ is the extension of $\lambda$ to $L(\mathscr{A})$; $L(\nu)$ is the *Loeb measure* (associated with $\nu$).

Strictly we should write $L(\mathscr{A}, \nu)$ instead of $L(\mathscr{A})$ to indicate that the completion of $\sigma(\mathscr{A})$ depends on $\nu$; however $L(\mathscr{A})$ has become accepted notation and it is always clear which $\nu$ is meant. The notation $L(\nu)$ for the Loeb measure has also become accepted; we find it convenient to introduce an alternative notation.

3.3. NOTATION. Write $\nu_L$ to denote the Loeb measure $L(\nu)$ associated with an internal measure $\nu$.

Important examples of Loeb spaces are the hyperfinite Loeb spaces discussed in subsection (3) below. Other natural examples include those derived from an internal space $^*\mathbf{X} = (^*X, {}^*\mathscr{F}, {}^*\mu)$ where $\mathbf{X}$ is standard; and more generally, if $\mathscr{M}_+$ denotes the collection of finitely additive measures on a measurable space $(X, \mathscr{F})$, then every $\nu \in {}^*\mathscr{M}_+$ is an internal f.a. measure on $^*X$ which has a corresponding Loeb measure $\nu_L$. Further Loeb measures can be obtained on carrier sets $Y$ that are not of the form $^*X$ for any $X$, as we shall see.

If $\nu$ is an internal finitely additive *signed* measure, then the Loeb construction can be carried out for $\nu$ by taking the decomposition $\nu = \nu^+ - \nu^-$ (where $\nu^+ = \nu \vee 0$ and $\nu^- = (-\nu) \vee 0$ in the lattice ordering of f.a. measures) and setting $\nu_L = \nu_L^+ - \nu_L^-$. It is necessary to assume that one of $\nu^+(Y)$ and $\nu^-(Y)$ is finite.

To simplify our exposition, *we shall restrict attention in the rest of this paper to finite Loeb measures*; that is, we shall assume that $\nu(Y)$ is finite and hence $\nu_L(Y) < \infty$. Most of the general results to be described extend with some modification to the unbounded case. Note that if $\nu_L(Y) < \infty$, the Loeb $\sigma$-algebra $L(\mathscr{A})$ is characterised as follows.

3.4. THEOREM. *Let $B \subseteq Y$; then $B$ is Loeb measurable if and only if*

$$\sup\{{}^\circ\nu(A) : A \subseteq B, A \in \mathscr{A}\} = \inf\{{}^\circ\nu(A) : B \subseteq A, A \in \mathscr{A}\},$$

*and then $\nu_L(B)$ is the common value of these expressions.*

On a historical note, we should mention that Robinson [**Rb**] came close to the construction of the Loeb measure for an internal space $\mathbf{Y} = {}^*\mathbf{X} = (^*X, {}^*\mathscr{F}, {}^*\mu)$, where $\mathbf{X}$ is standard. Writing $\mathscr{G}$ for the algebra $\{^*A : A \in \mathscr{F}\}$, Robinson showed (*without* using $\omega_1$-saturation) that the measure $^\circ(^*\mu)$ restricted to $\mathscr{G}$ has a unique extension $^*\mu_{\mathscr{S}}$ to $\sigma(\mathscr{G})$, which he called *$\mathscr{S}$-measure.* Of course $^*\mu_{\mathscr{S}}$ is the restriction of the Loeb measure $^*\mu_L$ to $\sigma(\mathscr{G})$. A nonstandard proof of Egoroff's theorem eluded Robinson, but one was provided by Henson and Wattenberg [**HW**] using $\mathscr{S}$-measure. It seems that the reason Robinson did not discover Loeb measure was his failure to exploit the $\omega_1$-saturation property.

### §3(2) *Loeb measurable functions and integration*

Much of the power of Loeb measure stems from the links between measurable and integrable functions on the Loeb space $L(\mathbf{Y}) = (Y, L(\mathscr{A}), \nu_L)$ and internal *measurable functions on $\mathbf{Y}$. This means that integration in $L(\mathbf{Y})$ can be obtained by *integration of suitable internal functions on $\mathbf{Y}$; if $\mathbf{Y}$ is hyperfinite this is simply summation.

The basic result regarding measurability in $L(\mathbf{Y})$ is the following.

3.5. LIFTING THEOREM (Loeb [**Lo**$_4$, **Lo**$_8$]). *Let $\mathbf{Y} = (Y, \mathscr{A}, \nu)$ be an internal finitely additive measure space with Loeb space $L(\mathbf{Y}) = (Y, L(\mathscr{A}), \nu_L)$. Let $f : \mathbf{Y} \to \overline{\mathbb{R}}$. Then the following are equivalent.*

(a) *$f$ is Loeb measurable (that is, $L(\mathscr{A})$-measurable);*

(b) *there is an internal $\mathscr{A}$-measurable function $F : Y \to {}^*\mathbb{R}$ such that ${}^\circ F(y) = f(y)$ for $\nu_L$-almost all $y \in Y$. ($\mathscr{A}$-measurable means that the sets $\{y : F(y) \leqslant \alpha\}$ and $\{y : F(y) \geqslant \alpha\}$ are in $\mathscr{A}$ for every $\alpha \in {}^*\mathbb{R}$.)*

*If $f$ is bounded, $F$ can be obtained with the same bound.*

A function $F$ satisfying (b) is called a *lifting* of $f$; so a lifting of $f$ is an $\mathscr{A}$-measurable function for which the following diagram commutes for a.a. $y \in Y$.

$$\begin{tikzcd} & {}^*\mathbb{R} \arrow[d, "\circ"] \\ Y \arrow[ur, "F"] \arrow[r, "f"'] & \overline{\mathbb{R}} \end{tikzcd}$$

In particular, if $F$ is $\mathscr{A}$-measurable then $F$ lifts the function ${}^\circ F$ which is Loeb measurable.

*Proof.* First consider an $\mathscr{A}$-measurable function $F : Y \to {}^*\mathbb{R}$; for any real $r$, we have

$$\{y : {}^\circ F(y) \leqslant r\} = \bigcap_{n \in \mathbb{N}} \{y : F(y) \leqslant r + 1/n\} \in \sigma(\mathscr{A}),$$

so ${}^\circ F$ is Loeb measurable. Thus, if $f(y) = {}^\circ F(y)$ a.s., $f$ is also Loeb measurable.

Conversely, suppose that $f$ is Loeb measurable; let $(q_n)_{n \in \mathbb{N}}$ enumerate the standard rationals $\mathbb{Q}$ and let $B_n = \{y : f(y) \leqslant q_n\}$. Find internal sets $A_n \in \mathscr{A}$ such that $\nu_L(A_n \Delta B_n) = 0$, and such that $A_n \subseteq A_m$ whenever $q_n \leqslant q_m$. Extend the sequence $(A_n)_{n \in \mathbb{N}}$ using denumerable comprehension, and by overflow find infinite $H$ such that

$$A_n \in \mathscr{A}, \text{ and } q_n \leqslant q_m \Rightarrow A_n \subseteq A_m \qquad \text{for all } n, m \leqslant H.$$

Since $\{q_1, \ldots, q_H\}$ is hyperfinite and $A_n \in \mathscr{A}$ for all $n \leqslant H$, define an internal $\mathscr{A}$-measurable $F : Y \to {}^*\mathbb{R}$ such that for every $n \leqslant H$, $F(y) \leqslant q_n \Leftrightarrow y \in A_n$.

Specifically, if $q_{i_1} < q_{i_2} < \ldots < q_{i_H}$ is the order of $q_1, \ldots, q_H$ in $^*\mathbb{Q}$, put

$$F(y) = \begin{cases} q_{i_1} & \text{if } y \in A_{i_1} \\ q_{i_j} & \text{if } y \in A_{i_j} \setminus A_{i_{j-1}} \quad (1 < j \leqslant H) \\ q_{i_H}+1 & \text{if } y \notin A_{i_H}. \end{cases}$$

Then the set $\bigcup_{n \in \mathbb{N}} (A_n \,\Delta\, B_n)$ is $\nu_L$-null, and for $y$ outside it

$$F(y) \leqslant q_n \Leftrightarrow f(y) \leqslant q_n \quad (n \in \mathbb{N}).$$

Hence ${}^\circ F(y) = f(y)$ for $\nu_L$-a.a. $y$.

REMARK. The lifting $F$ of $f$ produced above is actually a **simple function*; that is, $F$ has hyperfinite range.

The above theorem has been generalised so that we have a similar characterisation of Loeb measurable functions $f : Y \to M$, where $M$ is a complete separable metric space; see **[Lo$_4$]** and **[An$_2$]**.

Now we turn to the integration of bounded measurable functions.

3.6. THEOREM. *With the same preamble as Theorem* 3.5, *let $f$ be bounded Loeb measurable with bounded lifting $F$. Then*

$$\int f \, d\nu_L = {}^\circ\!\int F \, d\nu. \tag{3.7}$$

(The integral on the left is the integral in the sense of standard integration theory; that on the right is the internal integral of $F$, so strictly it should be written $^*\!\int F d\nu$.)

*Proof.* First observe that for bounded liftings $F_1$ and $F_2$ of $f$, since

$$|F_1(y) - F_2(y)| \leqslant 1/n \quad \text{for a.a. } y \quad (\text{for any } n \in \mathbb{N}), \quad \int F_1 \, d\nu \approx \int F_2 \, d\nu.$$

Now consider a simple Loeb measurable function $g \leqslant f$; let $g$ have finite range $\{r_1, \ldots, r_n\}$. Using the last part of Theorem 3.1 on the sets $g^{-1}(\{r_i\})$, obtain a lifting $G$ of $g$ with range $\{r_1, \ldots, r_n\}$; it is obvious that $\int g \, d\nu_L = {}^\circ\!\int G \, d\nu$, from the definition of the integrals in this case. Now take a bounded lifting $F_1$ of $f$ with $G \leqslant F_1$; then

$$\int g \, d\nu_L = {}^\circ\!\int G \, d\nu \leqslant {}^\circ\!\int F_1 \, d\nu = {}^\circ\!\int F \, d\nu$$

by the opening remark. Hence $\int f \, d\nu_L \leqslant {}^\circ\!\int F \, d\nu$. The reverse inequality follows by symmetry, considering $-f$ and $-F$.

For *unbounded* integrable $f$ it may not be possible to find a bounded lifting $F$ of $f$; however, with careful choice of $F$ according to the following definition the representation (3.7) of Loeb integration can be extended.

3.8. DEFINITION [**Lo**$_8$]. Let $F : Y \to {}^*\mathbb{R}$ (where $(Y, \mathscr{A}, \nu)$ is an internal finitely additive measure space). Then $F$ is *$\mathscr{S}$-integrable* if $F$ is $\mathscr{A}$-measurable and for every infinite $H \in {}^*\mathbb{N}$

$$\int_{|F| \geqslant H} |F| d\nu \approx 0 .$$

Clearly any finitely bounded $\mathscr{A}$-measurable function is $\mathscr{S}$-integrable.

3.9. THEOREM. *With the preamble of Theorem* 3.5, *the following are equivalent:*

(a) *$f$ is Loeb integrable;*

(b) *$f$ has an $\mathscr{S}$-integrable lifting $F : Y \to {}^*\mathbb{R}$.*

*If* (b) *holds, then* $\int f d\nu_L = {}^\circ\!\int F d\nu$.

*Proof.* Assume (a). We may suppose that $f \geqslant 0$ (for general $f$ consider $f = f^+ - f^-$). Let $F' \geqslant 0$ be a lifting of $f$; let $F_n = F' \wedge n$. Then by Theorem 3.6, and standard integration theory

$$^\circ\!\int F_n d\nu = \int (f \wedge n) d\nu_L \nearrow \int f d\nu_L .$$

Extend the sequence $(F_n)$ by denumerable comprehension, and by overflow find infinite $H$ such that $\int F_N d\nu \approx \int f d\nu_L$ for all infinite $N \leqslant H$. Taking $F = F_H$, it is clear that $F$ is $\mathscr{S}$-integrable and lifts $f$.

Conversely, if we are given an $\mathscr{S}$-integrable lifting $F \geqslant 0$ of $f$, use underflow to show that for every real $\varepsilon > 0$ there is finite $n$ such that, if $m \geqslant n$ and $m$ is finite,

$$\int_{F \geqslant m} F d\nu \leqslant \varepsilon .$$

Then using Theorem 3.6 we have, for finite $m \geqslant n$

$$\int (f \wedge m) d\nu_L \approx \int (F \wedge m) d\nu \leqslant \int F d\nu \leqslant \int (F \wedge m) d\nu + \varepsilon \approx \int (f \wedge m) d\nu_L + \varepsilon .$$

From this we see that ${}^\circ\!\int F d\nu$ is finite and is the limit of $\int (f \wedge m) d\nu_L$ as $m \to \infty$, as required.

Note that from the above proof, if $f \geqslant 0$ is integrable, then for *any* lifting $F \geqslant 0$ of $f$, $\int f d\nu_L \leqslant {}^\circ\!\int F d\nu$; the condition of $\mathscr{S}$-integrability is precisely what is needed to make these expressions equal; see Theorem 3.10(c) below. The definition of $\mathscr{S}$-integrability above was formulated by Loeb [**Lo**$_8$]; an earlier, equivalent definition (3.10(b), below) was given by Anderson [**An**$_1$].

3.10. THEOREM. *Let* $\mathbf{Y} = (Y, \mathscr{A}, \nu)$ *be an internal finitely additive measure space and let* $F : Y \to {}^*\mathbb{R}$ *be $\mathscr{A}$-measurable. Then the following are equivalent.*

(a) *$F$ is $\mathscr{S}$-integrable;*

(b) (i) $^\circ\int |F| dv < \infty$,

(ii) $\int_A F dv \approx 0$ *whenever* $A \in \mathscr{A}$ *with* $v(A) \approx 0$;

(c) $\int {}^\circ|F| dv_L = {}^\circ\int |F| dv < \infty$.

We leave a proof of this theorem as an exercise for the reader.

The condition (b)(ii) is a natural reflection of the continuity of the standard integral; if the space $L(\mathbf{Y})$ is atomless (i.e. for every $A \in \mathscr{A}$ with $^\circ v(A) > 0$ there is $B \subseteq A$, $B \in \mathscr{A}$ with $0 < {}^\circ v(B) < {}^\circ v(A)$) then (b)(ii) implies (b)(i) above.

In [**An**$_2$] and [**Lo**$_8$], $\mathscr{S}$-integrability was used to give a pleasant characterisation of uniform integrability in standard measure spaces.

3.11. THEOREM. *Let $\mathbf{X} = (\mathbf{X}, \mathscr{F}, \mu)$ be a standard measure space with $\mu$ finite. A collection of measurable functions $(f_i)_{i \in I}$ is uniformly integrable if and only if $f_i$ is $\mathscr{S}$-integrable for every $i \in {}^*I$.*

Using this we can obtain neat proofs of the following well known facts.

3.12. THEOREM. *Let $\mathbf{X} = (X, \mathscr{F}, \mu)$ be a measure space, with $\mu$ finite.*

(a) *If $(f_n)_{n \in \mathbb{N}}$ are uniformly integrable on $\mathbf{X}$ and $f_n \to 0$ in measure, then $\int f_n \to 0$.*

(b) *A uniformly integrable set of functions $(f_i)_{i \in I}$ is relatively weakly compact in the $(\mathscr{L}^1, \mathscr{L}^\infty)$-weak topology of $\mathscr{L}^1(\mathbf{X})$.*

*Proof.* (a) By Theorem 1.7(b) we have to show that $\int f_H d^*\mu \approx 0$ for all infinite $H$. Fix infinite $H$; for each finite $n$, by Theorem 1.7(b), we have $^*\mu(\{|f_H| \geqslant 1/n\}) \leqslant 1/n$. By overflow this holds for some infinite $N$, so $^\circ f_H = 0$ a.s. The $\mathscr{S}$-integrability of $f_H$ means that $\int f_H d^*\mu \approx 0$.

(b) Let $g = f_i$, where $i \in {}^*I$; by Theorem 1.24 and Proposition 1.27 we have to show that there is $h \in \mathscr{L}^1(\mathbf{X})$ with $\displaystyle\int_A h d\mu \approx \int_{^*A} g d^*\mu$ for all $A \in \mathscr{F}$. For $A \in \mathscr{F}$, define $\lambda(A) = {}^\circ\left(\displaystyle\int_{^*A} g d^*\mu\right)$; the $\mathscr{S}$-integrability of $g$ means that $\lambda$ is a signed measure on $\mathscr{F}$ and that $\lambda \ll \mu$. The Radon–Nikodym theorem gives $h \in \mathscr{L}^1(\mathbf{X})$ such that

$$\int_A h d\mu = \lambda(A) \approx \int_{^*A} g d^*\mu \quad \text{for all } A \in \mathscr{F},$$

as required.

In some recent work (reported in the forthcoming papers [**Lo**$_{11}$, **Lo**$_{12}$] Loeb has developed a Daniell-type extension of his theory of measure-construction. He begins with an internal vector lattice $L$ of $^*\mathbb{R}$-valued functions on an internal set $Y$, together with an internal positive linear functional $I$ on $L$. From this he obtains a standard vector lattice $L_1$ of real valued functions on $Y$ together with a *continuous* positive linear functional $J$ on $L_1$. The triple $(Y, L, I)$ is related to $(Y, L_1, J)$ in much the same way as an internal measure space is related to its Loeb space. Loeb obtains, moreover, a standard measure $\mu$ on $Y$ such that $J(f) = \int f d\mu$ for $f \in L_1$.

The simplest example of this functional approach is, of course, when $L$ is the lattice of $\mathscr{A}$-simple functions on an internal measure space $(Y, \mathscr{A}, \nu)$; then $L_1$ is the lattice of Loeb integrable functions and $\mu = \nu_L$.

Loeb's constructions in this new approach are achieved easily from first principles just as was possible with Loeb measure (see the proof of Theorem 3.1). An immediate application is the Riesz representation theorem.

## §3(3) *Hyperfinite Loeb spaces*

Some of the most important Loeb spaces are those where the underlying set $Y$ is hyperfinite. An internal algebra $\mathscr{A} \subseteq {}^*\mathscr{P}(Y)$ is necessarily hyperfinite; any measure $\nu$ on $\mathscr{A}$ could be extended to the whole of ${}^*\mathscr{P}(Y)$ (by transfer of the appropriate standard theorem). It is usual when $Y$ is hyperfinite to assume that $\mathscr{A} = {}^*\mathscr{P}(Y)$.

If $\mathbf{Y} = \big(Y, {}^*\mathscr{P}(Y), \nu\big)$ with $Y$ hyperfinite, then $\nu$ is given by a weighting function $\Delta\nu : Y \to {}^*\mathbb{R}$, where

$$\Delta\nu(y) = \nu(\{y\}) \quad \text{for } y \in Y,$$

and the measure $\nu$ is a weighted counting measure:

$$\nu(A) = \sum_{y \in A} \Delta\nu(y) \quad \text{for internal } A \subseteq Y.$$

The simplest example of this is the internal counting probability measure on $Y$, usually denoted $\bar{P}_Y$ or just $\bar{P}$, given by

$$\bar{P}(A) = |A|/|Y| \quad \text{for } A \subseteq Y,\ A \text{ internal}.$$

Then $\Delta\bar{P}(y) = |Y|^{-1}$ for all $y$. The associated Loeb space is a probability space, the *uniform Loeb probability space* on $Y$; the uniform Loeb probability measure $L(\bar{P}_Y)$ is denoted $P_Y$ or just $P$ if the context makes it clear. So for internal $A \subseteq Y$ we have $P(A) = {}^\circ(|A|/|Y|)$. It is convenient to denote $L\big({}^*\mathscr{P}(Y)\big)$ by $L(Y)$.

A simple example of a non-Loeb measurable set is obtained as follows. Take $Y = \{1, 2, \ldots, H\}$, where $H$ is infinite. For $M \in Y$ let

$$\text{galaxy}(M) = \{N \in Y : M - N \text{ is finite } (\operatorname{mod} H)\}.$$

Let $D \subseteq Y$ be a set consisting of one point chosen from each galaxy in $Y$. Then $Y = \bigcup_{n \in \mathbb{Z}} (D + n)$, which is a disjoint union; if $D$ were Loeb measurable we would have $P(D) = P(D + n)$ for all $n$, which is impossible.

In later sections we will see how uniform Loeb measure can be used to give pleasant constructions of standard measures such as Lebesgue measure and Wiener measure.

The lifting theorem (Theorem 3.5) for a hyperfinite Loeb space $L\big(Y, {}^*\mathscr{P}(Y), \nu\big)$ is slightly simplified because an internal function $F : Y \to {}^*\mathbb{R}$ is automatically

$*\mathscr{P}(Y)$-measurable. Integration on $\mathbf{Y}$ is of course summation, so if $F$ is an $\mathscr{S}$-integrable lifting of a measurable function $f: Y \to \bar{\mathbb{R}}$ we have $\int f d\nu_L = {}^\circ\sum_{y \in Y} F(y)\Delta\nu(y)$. In the case that $\nu = \bar{P}_Y$, it is convenient to define $\Delta y = |Y|^{-1}$ $(= \Delta\nu(y))$, so that $\int f dP = {}^\circ\sum_{y \in Y} F(y)\Delta y$. Note that the Loeb space $(Y, L(Y), P)$ is atomic (if $Y$ is infinite); thus, by Theorem 3.10(b) and the remark following it, an internal function $F: Y \to {}^*\mathbb{R}$ is $\mathscr{S}$-integrable if and only if $\sum_{y \in A} F(y)\Delta y \approx 0$ whenever $A \subseteq Y$ is internal and $\bar{P}(A) \approx 0$.

To see the need for $\mathscr{S}$-integrability, pick $y_0 \in Y$ and define $F(y) = 0$ $(y \neq y_0)$; $F(y_0) = |Y|$. Then $F$ lifts $f = 0$, but $\sum F(y)\Delta y = 1$ whereas $\int f dP = 0$.

### §3(4) *Products of Loeb spaces*

If $\mathbf{Y}_1 = (Y_1, \mathscr{A}_1, \nu_1)$ and $\mathbf{Y}_2 = (Y_2, \mathscr{A}_2, \nu_2)$ are two internal finitely additive measure spaces, we can construct from their Loeb spaces the standard product space

$$L(\mathbf{Y}_1) \times L(\mathbf{Y}_2) = \big(Y_1 \times Y_2, L(\mathscr{A}_1) \times L(\mathscr{A}_2), (\nu_1)_L \times (\nu_2)_L\big).$$

Let $\overline{L(\mathscr{A}_1) \times L(\mathscr{A}_2)}$ denote the completion of the product $\sigma$-algebra with respect to the product measure.

There is a second natural standard product that is obtained by reversing the order of the Loeb and product operations. We have the *internal* product space

$$\mathbf{Y}_1 \times \mathbf{Y}_2 = (Y_1 \times Y_2, \mathscr{A}_1 \times \mathscr{A}_2, \nu_1 \times \nu_2)$$

where $\mathscr{A}_1 \times \mathscr{A}_2$ is the internal product algebra generated by $\mathscr{A}_1$ and $\mathscr{A}_2$; this space has corresponding Loeb space

$$L(\mathbf{Y}_1 \times \mathbf{Y}_2) = \big(Y_1 \times Y_2, L(\mathscr{A}_1 \times \mathscr{A}_2), (\nu_1 \times \nu_2)_L\big).$$

In [$\mathbf{An}_1$] Anderson showed that $\overline{L(\mathscr{A}_1) \times L(\mathscr{A}_2)} \subseteq L(\mathscr{A}_1 \times \mathscr{A}_2)$, and that $(\nu_1 \times \nu_2)_L$ agrees with $(\nu_1)_L \times (\nu_2)_L$ on the smaller $\sigma$-algebra; however the following example shows that in general the inclusion is proper.

3.13. Example (Hoover). Let $Y$ be infinite, hyperfinite and let $Z = {}^*\mathscr{P}(Y)$; so $Z$ is hyperfinite and $|Z| = 2^{|Y|}$; take the uniform Loeb measures $P_Y$ and $P_Z$ on $Y$ and $Z$. Then the set $D = \{(y, z) : y \in z\} \subseteq Y \times Z$, being internal, is in $L(Y \times Z)$; clearly $P_{Y \times Z}(D) = \frac{1}{2}$. So, if $D$ were in $\overline{L(Y) \times L(Z)}$ there would be internal $A \times B \subseteq D$ with $P_Y(A) > 0$ and $P_Z(B) > 0$. But $A \times B \subseteq D$ implies that $B \subseteq B_1 = \{z : z$ internal, $A \subseteq z \subseteq Y\}$, and simple calculation shows that $P_Z(B_1) = {}^\circ(2^{-|A|}) = 0$.

In view of the above example, the following Fubini theorem is a significant extension of the standard Fubini theorem for the product $L(\mathbf{Y}_1) \times L(\mathbf{Y}_2)$ of two Loeb spaces.

3.14. THEOREM (Keisler's Fubini Theorem). *Let* $\mathbf{Y}_i = (Y_i, \mathscr{A}_i, \nu_i)$ $(i = 1, 2)$ *be internal finitely additive measure spaces and let* $f : Y_1 \times Y_2 \to \overline{\mathbb{R}}$ *be* $L(\mathscr{A}_1 \times \mathscr{A}_2)$ *measurable. Then*

(a) $f(\cdot, y_2)$ is $L(\mathscr{A}_1)$*-measurable for* $(\nu_2)_L$-a.a. $y_2 \in Y_2$;

(b) *if* $f$ *is integrable, then*

 (i) $f(\cdot, y_2)$ *is integrable over* $Y_1$ *for* a.a. $y_2$,

 (ii) *the function* $g(y_2) = \int f(y_1, y_2) d(\nu_1)_L$ *is integrable over* $Y_2$,

 (iii) $\int g(y_2) d(\nu_2)_L = \int f(y_1, y_2) d(\nu_1 \times \nu_2)_L$.

This theorem was proved for hyperfinite spaces in [**Ke**$_1$]; a simpler proof is given in [**Ke**$_3$]. The crux of the matter is to show that if $A \in L(\mathscr{A}_1 \times \mathscr{A}_2)$ is null, then for a.a. $y_2$, the section $A_{y_2}$ is null in $L(\mathscr{A}_1)$. The rest follows by taking a lifting $F$ of $f$ with respect to $(\nu_1 \times \nu_2)_L$.

## §4 REPRESENTATIONS OF MEASURES BY LOEB MEASURES

### §4(1) *Lebesgue measure*

Uniform Loeb measure on a hyperfinite set gives an easy construction of Lebesgue measure, which we now describe. For convenience we restrict to $[0, 1]$. Fix infinite $H$ in ${}^*\mathbb{N}$ and let $\Delta t = H^{-1}$. Let $T$ be the hyperfinite set $T = \{0, \Delta t, 2\Delta t, \ldots, H\Delta t\}$. Then $T$ gives a representation of $[0, 1]$ via the standard part map $\mathrm{st} = {}^\circ : T \to [0, 1]$. Here and elsewhere we will use the convention that sanserif symbols $\mathsf{s}, \mathsf{t}, \mathsf{u}, \ldots$ denote elements of $T$; some authors use underlined variables for elements of $T$, which is convenient in manuscripts. In the following $P$ denotes $P_T$.

4.1. THEOREM. *Let* $A \subseteq [0, 1]$; *the following are equivalent:*

(a) *$A$ is Lebesgue measurable,*

(b) $\mathrm{st}^{-1}(A)$ *is a Loeb measurable subset of* $T$.

*If* (b) *holds, then* $m(A) = P(\mathrm{st}^{-1}(A))$.

*Proof.* Note that st here is the standard part map *restricted to* $T$, so $\mathrm{st}^{-1}(A) = \{\mathsf{t} \in T : \mathrm{st}(\mathsf{t}) \in A\}$. Consider first an interval $[a, b] \subseteq [0, 1]$; then

$$\mathrm{st}^{-1}([a, b]) = \bigcap_{n \in \mathbb{N}} ({}^*(a - 1/n, b + 1/n) \cap T)$$

which is in $L(T)$. It is easy to check that $P(\mathrm{st}^{-1}([a, b])) = b - a$. Now $\mathrm{st}^{-1}$ preserves differences and countable unions, so $\mathrm{st}^{-1}(B) \in L(T)$ for all Borel $B$. This extends to all Lebesgue sets $A$ since $L(T)$ is complete.

For the converse, suppose that $A \subseteq [0, 1]$ and $\mathrm{st}^{-1}(A) \in L(T)$; take $D \subseteq \mathrm{st}^{-1}(A)$, $D$ internal with $P(D) \geqslant P(\mathrm{st}^{-1}(A)) - \varepsilon$. Now $\mathrm{st}(D) = C$, say, is closed, by Theorem 1.26; and $C \subseteq A$. Moreover, $D \subseteq \mathrm{st}^{-1}(C) \subseteq \mathrm{st}^{-1}(A)$. Using the first part of the proof above we have

$$m(C) = P(\mathrm{st}^{-1}(C)) \geqslant P(D) \geqslant P(\mathrm{st}^{-1}(A)) - \varepsilon,$$

which shows that $m_{\text{inner}}(A) \geqslant P(\text{st}^{-1}(A))$. A similar argument applied to $A^c$ gives $m_{\text{outer}}(A) \leqslant P(\text{st}^{-1}(A))$, so the theorem is proved.

Given the framework of the nonstandard universe and the uniform Loeb measure on hyperfinite sets, the above may be taken as a construction of Lebesgue measure. Lebesgue measurable functions are now represented by extending the idea of a lifting.

4.2. DEFINITION. Let $f : [0, 1] \to \overline{\mathbb{R}}$ and let $F : T \to {}^*\mathbb{R}$; then $F$ is a *lifting* of $f$ if $F$ is internal and

$$°F(\mathrm{t}) = f(°\mathrm{t}) \text{ for } P\text{-almost all } \mathrm{t} \in T.$$

That is, $F$ is a lifting, in the sense of Theorem 3.5, of the function $\hat{f} : T \to \overline{\mathbb{R}}$ given by $\hat{f}(\mathrm{t}) = f(°\mathrm{t})$. A lifting of $f$ is an internal function that makes the following diagram commute for almost all $\mathrm{t} \in T$:

$$\begin{array}{ccc} T & \xrightarrow{\quad F \quad} & {}^*\mathbb{R} \\ {\scriptstyle\circ}\big\downarrow & & \big\downarrow{\scriptstyle\circ} \\ [0,1] & \xrightarrow[\quad f \quad]{} & \overline{\mathbb{R}} \end{array}$$

We obtain the expected characterisation of Lebesgue measurable functions and the Lebesgue integral by combining Theorem 4.1 with Theorems 3.5 and 3.9.

4.3. THEOREM. *Let* $f : [0, 1] \to \overline{\mathbb{R}}$,

(a) *The following are equivalent:*

(i) $f$ *is Lebesgue measurable,*

(ii) $\hat{f}$ *is Loeb measurable,*

(iii) $f$ *has a lifting.*

(b) $f$ *is Lebesgue integrable if and only if* $f$ *has an* $\mathcal{S}$*-integrable lifting; if* $F$ *is an* $\mathcal{S}$*-integrable lifting of* $f$ *then* $\int f\,dm = {}^\circ\sum_{\mathrm{t}\in T} F(\mathrm{t})\Delta t$.

In [**Rd**$_1$] Rodenhausen showed that a specific $\mathcal{S}$-integrable lifting for $f \in \mathcal{L}^1[0, 1]$ is obtained by setting

$$F(\mathrm{t}) = \Delta t^{-1} \int_{\mathrm{t}}^{\mathrm{t}+\Delta t} {}^*f\,d{}^*m \quad (\mathrm{t} \in T).$$

It is clear that not every internal function $G : T \to {}^*\mathbb{R}$ is a lifting of some function $g : [0, 1] \to \overline{\mathbb{R}}$. Rodenhausen obtained a variational principle that characterises those internal $G$ that are liftings. He also considered liftings of functions $f : [0, 1] \to \overline{\mathbb{R}}$ by ${}^*$measurable functions $F : {}^*[0, 1] \to {}^*\mathbb{R}$ (this time with respect to $({}^*m)_L$ where $m$ is Lebesgue measure) and found the following characterisation of nearstandardness in $\mathcal{L}^1[0, 1]$.

4.4. THEOREM [$\mathbf{Rd}_1$]. *Let $f:[0,1] \to \bar{\mathbb{R}}$ and $F: {}^*[0,1] \to {}^*\mathbb{R}$; then the following are equivalent:*

(a) *$F$ is nearstandard in ${}^*\mathscr{L}^1[0,1]$ and ${}^\circ F = f$ (with respect to the $\mathscr{L}^1$-norm),*

(b) *$F$ is an $\mathscr{S}$-integrable lifting of $f$; that is, $F \in {}^*\mathscr{L}^1[0,1]$, $F$ is $\mathscr{S}$-integrable and ${}^\circ F(x) = f({}^\circ x)$ for $({}^*m)_L$-almost all $x \in {}^*[0,1]$.*

§4(2) *Measures on Hausdorff spaces*

The representation of Lebesgue measure by a Loeb measure on a hyperfinite set was generalised by Anderson [$\mathbf{An}_2$] to an arbitrary Radon measure on a Hausdorff space. First recall the definition of a Radon measure.

DEFINITION. Let $(X, \tau)$ be a Hausdorff space and let $\mathscr{B}$ be the Borel $\sigma$-algebra on $X$. A finite Borel measure $\mu$ is *Radon* if for every $B \in \mathscr{B}$

$$\mu(B) = \sup\{\mu(K): K \subseteq B, K \text{ compact}\}.$$

The space $(X, \mathscr{B}, \mu)$ is called a *Radon measure space.*

Recall that for a Hausdorff space $X$ there is a standard part map $\mathrm{st}: \mathrm{ns}({}^*X) \to X$; see Definition 1.22. Anderson showed first that the completion of a Radon space is representable via the standard part map as a Loeb measure on ${}^*X$.

4.5. THEOREM [$\mathbf{An}_2$]. *Let $\mathbf{X} = (X, \mathscr{B}, \mu)$ be a Radon space, and let $\mathscr{C}$ be the $\mu$-completion of $\mathscr{B}$. Then $(X, \mathscr{C}, \mu)$ is represented on the Loeb space $({}^*X, L({}^*\mathscr{C}), ({}^*\mu)_L)$ as follows. For any $B \subseteq X$ the following are equivalent:*

(a) *$B \in \mathscr{C}$;*
(b) *$\mathrm{st}^{-1}(B) \in L({}^*\mathscr{C})$.*

*If $B \in \mathscr{C}$, then $\mu(B) = ({}^*\mu)_L(\mathrm{st}^{-1}(B))$; in particular $({}^*\mu)_L(\mathrm{ns}({}^*X)) = \mu(X)$.*

The proof of this theorem is a straightforward generalisation of the ideas in the proof of Theorem 4.1; the implication (b) $\Rightarrow$ (a) requires $\kappa$-saturation for suitable $\kappa$, so that Theorem 1.26 may be employed.

Next we mention Anderson's analogue of Lusin's theorem (the latter says, roughly, that a measurable function on a Hausdorff space is "almost continuous").

4.6. THEOREM [$\mathbf{An}_2$]. *Let $(X, \mathscr{B}, \mu)$ be a Radon space and let $f: X \to \bar{\mathbb{R}}$ be measurable. Then ${}^*f(x) \approx f({}^\circ x)$ for $({}^*\mu)_L$-almost all $x \in {}^*X$.*

*Thus there is a set $Y \subseteq {}^*X$ of full Loeb measure such that for $y_1, y_2 \in Y$, if $y_1 \approx y_2$ then ${}^*f(y_1) \approx {}^*f(y_2)$.*

In [$\mathbf{An}_2$] Theorem 4.5 was actually established with the internal algebra ${}^*\mathscr{C}$ replaced by any internal algebra $\mathscr{A} \subseteq {}^*\mathscr{C}$ such that for each open set $U \subseteq X$ either ${}^*U \in \mathscr{A}$ or $\mathrm{st}^{-1}(U) \in L(\mathscr{A})$. We now sketch the way this was used by Anderson to obtain a *hyperfinite* Loeb space representation of $X$.

Take $\mathscr{A} \subseteq {}^*\mathscr{C}$ such that $\mathscr{A} \supseteq \{{}^*C : C \in \mathscr{C}\}$ and $\mathscr{A}$ is hyperfinite, using Theorem 1.14. Let $\{A_1, \ldots, A_H\}$ be the partition of ${}^*X$ given by $\mathscr{A}$ and choose $x_i \in A_i$ for each $i$. Let $Y = \{x_1, \ldots, x_H\}$ and define an internal finitely additive measure $\nu$ on $Y$ by

$\Delta\nu(x_i) = {}^*\mu(A_i)$. Then $(Y, {}^*\mathscr{P}(Y), \nu)$ is essentially a quotient space of $({}^*X, \mathscr{A}, {}^*\mu)$ and using Theorem 4.5 we obtain the next theorem.

4.7. THEOREM [**An**$_2$]. *Let $(X, \mathscr{B}, \mu)$ be Radon, and let $Y \subseteq {}^*X$ be a hyperfinite set such as described above, and let $\nu$ be the corresponding internal measure on $Y$. Then for $B \subseteq X$, $B \in \mathscr{C}$ if and only if $\mathrm{st}_Y^{-1}(B)$ is Loeb measurable, where $\mathscr{C}$ is the completion of $\mathscr{B}$, and $\mathrm{st}_Y$ is the restriction of $\mathrm{st} : \mathrm{ns}({}^*X) \to X$ to $Y$. If $B \in \mathscr{C}$, then $\mu(B) = \nu_L(\mathrm{st}_Y^{-1}(B))$.*

The internal space $(Y, {}^*\mathscr{P}(Y), \nu)$ can be modified so as to represent **X** by a *uniform* Loeb space. For simplicity assume that **X** is a probability space; choose infinite $K$ with $H/K \approx 0$; choose internal $(n_y)_{y \in Y}$ from ${}^*\mathbb{N}$ so that $n_y/K \approx \Delta\nu(y)$, and $\sum_{y \in Y} n_y = K$. Working internally, choose internal disjoint hyperfinite sets $Z_y$ with $|Z_y| = n_y$, and set $Z = \bigcup_{y \in Y} Z_y$. Then $|Z| = K$; finally define $\pi : Z \to X$ by $\pi(z) = \mathrm{st}(y)$ if $z \in Z_y$ and $y \in \mathrm{ns}({}^*X) \cap Y$. (Of course $\pi$ is not necessarily defined on all of $Z$.) The following result has been obtained.

4.8. THEOREM. *Let $\mathbf{X} = (X, \mathscr{B}, \mu)$ be a Radon probability space and let $\mathscr{C}$ be the completion of $\mathscr{B}$. Let $Z$ and $\pi : Z \to X$ be defined as above (so $Z$ is hyperfinite). Then for $B \subseteq X$, $B \in \mathscr{C}$ if and only if $\pi^{-1}(B) \in L(Z)$. If $B \in \mathscr{C}$, then $\mu(B) = P_Z(\pi^{-1}(B))$.*

Note that in this representation theorem the infinite hyperfinite set $Z$ could be chosen arbitrarily in advance. We would simply have to choose the hyperfinite algebra $\mathscr{A}$ above so that $|\mathscr{A}|/|Z| \approx 0$.

If $(X, \mathscr{B}, \mu)$ is a general Borel measure on a Hausdorff space $(X, \tau)$ we can still form the Loeb space $({}^*X, L({}^*\mathscr{B}), {}^*\mu_L)$, but now the mapping $\mathrm{st}^{-1}$ may not be measure preserving. However, restricting $\mu$ to the Baire subsets of $X$ we do still get some positive results. The *Baire $\sigma$-algebra* $\mathscr{B}_0$ is the smallest $\sigma$-subalgebra making all functions in $C(X)$ measurable; i.e., in the usual terminology, $\mathscr{B}_0 = \sigma(C(X))$. Note that for a metric space $\mathscr{B}_0 = \mathscr{B}$. In [**He**$_2$] Henson showed that if $X$ is compact then it is precisely the Baire sets $A \subseteq X$ that satisfy $\mathrm{st}^{-1}(A) \in \sigma({}^*\mathscr{B})$. It was noted by Anderson [**An**$_2$] that a finite Baire measure $(X, \mathscr{B}_0, \mu)$ is represented on $({}^*X, L({}^*\mathscr{B}_0), {}^*\mu_L)$ by $\mathrm{st}^{-1}$ at least when $\mu$ is tight, which is the case if $X$ is compact or $\sigma$-compact; moreover, in this case the hyperfinite representation Theorems 4.7 and 4.8 can be obtained for Baire measures.

## §4(3) *Representation of general measures*

For a general measure space $\mathbf{X} = (X, \mathscr{F}, \mu)$ (with no topology assumed) it is well known that $\mathscr{L}^\infty(\mathbf{X}) \cong C(\hat{X})$ where $(\hat{X}, \hat{\tau})$ is the Stone space of **X**, and there is a Baire measure $\hat{\mu}$ on $\hat{X}$ with the property $\int f\,d\mu = \int \hat{f}\,d\hat{\mu}$ where $\hat{f} \in C(\hat{X})$ corresponds to $f \in \mathscr{L}^\infty(X)$. In [**An**$_2$] Anderson simultaneously obtained this representation together with a Loeb space representation of $\hat{\mu}$ as follows.

Let $\hat{X} = {}^*X/\sim$, where $x_1 \sim x_2$ if and only if ${}^*f(x_1) \approx {}^*f(x_2)$ for all $f \in \mathscr{L}^\infty(\mathbf{X})$. Let $\pi : {}^*X \to \hat{X}$ denote the quotient map sending $x$ to $x/\sim$; now for each $f \in \mathscr{L}^\infty(X)$ define $\hat{f}$ on $\hat{X}$ by $\hat{f}(\pi(x)) = {}^\circ({}^*f(x))$. The topology $\hat{\tau}$ is defined on $\hat{X}$ as the weakest topology making each $\hat{f}$ continuous. Finally the Baire measure $\hat{\mu}$ is defined on $\hat{X}$ by $\hat{\mu}(B) = ({}^*\mu)_L(\pi^{-1}(B))$.

Note that the above construction of the Stone space $\hat{X}$ is essentially the same as that given by Loeb [$\mathbf{Lo_2}$], which we described in §2(2). As described there, it is possible to replace $*X$ by a hyperfinite subset of $*X$.

An alternative approach to the representation of general measures on Loeb spaces was taken by Lindstrøm [$\mathbf{Li_1}$], who established the next result.

4.9. THEOREM. *Let* $(X, \mathscr{F}, \mu)$ *be a probability space and let* $(Z, L(Z), P_Z)$ *be any infinite hyperfinite Loeb space with the uniform Loeb probability measure* $P_Z$. *Then there is a* $\sigma$*-algebra* $\mathscr{G} \subseteq L(Z)$ *and a surjective measure preserving* $\sigma$*-homomorphism* $\theta : \mathscr{G} \to \mathscr{F}$; *that is,* $P_Z(A) = \mu(\theta(A))$ *for all* $A \in \mathscr{G}$.

Lindstrøm calls a representation such as this a *weak Loeb space representation*, compared to those given in Theorems 4.5, 4.7, and 4.8 above, in which representation is by pointwise maps between the spaces involved. In [$\mathbf{Li_1}$] weak representations were obtained for adapted probability spaces, suitable for carrying a full theory of stochastic analysis; see §7.

## §5 CONSTRUCTION OF MEASURES

The Loeb construction combined with the map $\mathrm{st}^{-1}$ has been shown by Anderson, Lindstrøm, Loeb and others to be an efficient way to construct measures on standard spaces—for example the construction of extensions of finitely additive measures on a given algebra, or the construction of limit measures. In this section we discuss a sample of results of this kind.

### §5(1) *Extension of measures*

In [$\mathbf{Li_2}$] Lindstrom gave an intuitive Loeb measure proof of the following 'classical' theorem due to Prohorov.

5.1. THEOREM. *Let* $((X_i, \tau_i); \pi_{ij})_{i,j \in I}$ *be a projective system of Hausdorff spaces with projective limit* $X$ *and projections* $\pi_i : X \to X_i$. *Suppose that* $(\mu_i)_{i \in I}$ *is a cylindrical probability measure on this system; that is, each* $\mu_i$ *is a Radon probability measure on* $(X_i, \tau_i)$ *and* $\mu_i(A) = \mu_j(\pi_{ij}^{-1}(A))$ *for Borel* $A \subseteq X_i$. *Then the following are equivalent:*

(a) *there is a Radon probability* $\mu$ *on* $X$ *such that* $\mu_i(A) = \mu(\pi_i^{-1}(A))$ *for each* $A \subseteq X_i$, *and for all* $i$;

(b) *for each* $\varepsilon > 0$ *there is a compact set* $K_\varepsilon \subseteq X$ *such that* $\mu_i(\pi_i(K_\varepsilon)) \geqslant 1 - \varepsilon$ *for all* $i \in I$.

*Proof* (sketch). That (a) implies (b) is almost immediate. Conversely, to construct $\mu$ given condition (b), take $\omega \in *I$ with $\omega > i$, for all $i \in I$ (using the enlargement axiom). The internal measure $\mu_\omega$ on $X_\omega$ yields the Loeb measure $(\mu_\omega)_L$. We have the projection $\pi_\omega : *X \to X_\omega$ and the standard part map $\mathrm{st} : *X \to X$, so we define a measure $\mu$ on $X$ by $\mu(B) = (\mu_\omega)_L(\pi_\omega(\mathrm{st}^{-1}(B)))$.

The rest of the proof now consists in showing that $\mu(B)$ is defined for all Borel $B$, and checking that $\mu$ has the required properties.

Similar ideas were used in [$\mathbf{Li_2}$] to prove other extension theorems, including a result of Gross concerning the extension of a cylindrical measure on the projective system of finite dimensional subspaces of a separable Hilbert space. This, and Prohorov's theorem, are obtained in [**AFHL**] as applications of the following strong extension theorem, again proved using Loeb measure.

5.2. THEOREM. *Let $(X, \tau)$ be a Hausdorff space, with $\tau_0 \subseteq \tau$ a basis for the topology. Let $\mathscr{D}$ be the algebra generated by $\tau_0$, and suppose that $\mu_0$ is a finitely additive regular probability measure on $\mathscr{D}$; assume that for each real $\varepsilon > 0$ there is a compact set $K_\varepsilon$ with $\beta_{K_\varepsilon} > 1-\varepsilon$, where $\beta_K = \inf\{\mu_0(D) : K \subseteq D \in \mathscr{D}\}$. Then there is a unique Radon probability measure $\mu$ on $X$ extending $\mu_0$; moreover, $\mu(K) = \beta_K$ for each compact $K \subseteq X$. (The f.a. measure $\mu_0$ is* regular *if $\mu_0(D) = \sup\{\mu_0(F) : D \supseteq F \in \mathscr{D}$, $F$ closed$\}$.)*

*Proof.* The proof of this theorem requires $\kappa$-saturation, where $\kappa > \operatorname{card}(\tau_0)$. We sketch the non-routine parts of the proof, as given in [**AFHL**].

The Loeb construction gives a probability space $({}^*X, L({}^*\mathscr{D}), ({}^*\mu_0)_L)$ which is used to define a probability measure on $X$ via $\mathrm{st}^{-1}$. Specifically, let $\mathscr{B}$ be the Borel $\sigma$-algebra on $X$ and define $\mathscr{F}$ and $\mu$ by

$$\mathscr{F} = \{B \in \mathscr{B} : \mathrm{st}^{-1}(B) \in L({}^*\mathscr{D})\}$$

$$\mu(B) = ({}^*\mu_0)_L(\mathrm{st}^{-1}(B)) \quad \text{for } B \in \mathscr{F}\,.$$

Check first that $\mathscr{F} = \mathscr{B}$. Begin with compact $K \subseteq X$; show that

$$\mathrm{st}^{-1}(K) = \bigcap\{{}^*U : K \subseteq U \in \mathscr{D},\, U \text{ open}\}\,.$$

Since $\mu_0$ is regular,

$$\beta_K = \inf\{\mu_0(U) : K \subseteq U \in \mathscr{D},\, U \text{ open}\}\,.$$

Now use $\kappa$-saturation to find $A \in {}^*\mathscr{D}$ such that $A \subseteq {}^*U$ for each open $U \in \mathscr{D}$ with $U \supseteq K$, and such that ${}^*\mu_0(A) \geqslant \beta_K$. Then $A \subseteq \mathrm{st}^{-1}(K)$, and so $\mathrm{st}^{-1}(K) \in L({}^*\mathscr{D})$ with $({}^*\mu_0)_L(\mathrm{st}^{-1}(K)) = \beta_K$. Hence $K \in \mathscr{F}$ and $\mu(K) = \beta_K$.

Let $X_1 = \bigcup_{n \in \mathbb{N}} K_{1/n}$; the condition on $K_{1/n}$ means that $\mathrm{st}^{-1}(X_1)$ has full Loeb measure, hence $\mathrm{st}^{-1}(X) \in L({}^*\mathscr{D})$. It is now easy to see that $F \in \mathscr{F}$ for each closed set $F$, so $\mathscr{F} = \mathscr{B}$. To show that $\mu$ is regular (and hence Radon) proceed as in the proof of Theorem 4.1 (use the fact that if $A$ is internal and $A \subseteq \mathrm{st}^{-1}(B)$ for Borel $B$, then $\mathrm{st}(A)$ is closed and $\mathrm{st}(A) \subseteq B$).

To check that $\mu$ extends $\mu_0$, let $D \in \mathscr{D}$, and take closed and open approximations $F, U \in \mathscr{D}$ with $F \subseteq D \subseteq U$. Then note that

$${}^*F \cap \mathrm{st}^{-1}(X) \subseteq \mathrm{st}^{-1}(F) \subseteq \mathrm{st}^{-1}(D) \subseteq \mathrm{st}^{-1}(U) \subseteq {}^*U$$

and $({}^*\mu_0)_L({}^*F \cap \mathrm{st}^{-1}(X)) = \mu_0(F)$ and $({}^*\mu_0)_L({}^*U) = \mu_0(U)$.

For uniqueness, suppose that $\bar{\mu}$ is another Radon extension of $\mu_0$; clearly $\bar{\mu}(K) \leqslant \beta_K = \mu(K)$ for compact $K$, so $\bar{\mu}(B) \leqslant \mu(B)$ for all Borel $B$; for the opposite inequality take complements.

This theorem is used in [**AFHL**] to obtain Wiener measure on $C[0, 1]$. Here the topology $\tau$ on $C[0, 1]$ is generated by cylinder sets of the form

$$D = \bigcap_{i=1}^{n} \{x \in C[0, 1] : x_{t_i} \in U_i\}$$

where $0 < t_1 < t_2 < \ldots < t_n \leqslant 1$ and $U_1, \ldots, U_n$ are open. Wiener measure is the unique Borel probability measure $\mu_w$ on $(C[0,1], \tau)$ such that

$$\mu_w(D) = \int_{U_n} \cdots \int_{U_1} \prod_{i=1}^{n} \phi(y_i - y_{i-1}; t_i - t_{i-1}) dy_1 \ldots dy_n \tag{5.3}$$

(where $t_0 = y_0 = 0$) where $\phi(y; \sigma^2) = (2\pi\sigma^2)^{-1/2} \exp(-y^2/2\sigma^2)$ is the density of the normal distribution $N(0, \sigma^2)$.

To use Theorem 5.2 to show the existence of Wiener measure, first check that (5.3) defines a finitely additive regular measure on the algebra $\mathscr{D}$ generated by the open cylinder sets. Then consider the sets $K_n^\alpha$ where

$$K_n^\alpha = \{x \in C[0,1] : 0 < t - s \leqslant 1/n \Rightarrow (x_t - x_s)^2 \leqslant \alpha(s-t)\log(t-s)\}.$$

Use the criterion 1.24 to show that $K_n^\alpha$ is compact (even in the stronger topology of uniform convergence). Now show by fairly simple calculation that if $\alpha > 2$ then $\beta_{K_n^\alpha} \to 1$ as $n \to \infty$; the extension now given by Theorem 5.2 is Wiener measure.

It is shown further in **[AFHL]** that for $\alpha < 2$, $\beta_{K_n^\alpha} \to 0$ as $n \to \infty$, and so for $\mu_w$-almost all $x \in C[0,1]$

$$\overline{\lim_{\delta \to 0}} \sup_{t \in [0,1]} \frac{x(t+\delta) - x(t)}{(2\delta \log(\delta^{-1}))^{1/2}} = 1,$$

which is Lévy's famous modulus of continuity result.

### §5(2) *Weak limits of measures*

In **[AR]** and **[Lo$_7$]** Anderson–Rashid and Loeb show how Loeb measures can be used to obtain weak limits of measures on a topological space via the standard part map. To illustrate their results, fix a compact Hausdorff space $(X, \tau)$ and let $\mathscr{B}_0$ denote the Baire $\sigma$-algebra on $X$.

Let $\mathscr{M}(X)$ be the set of all finite signed Baire measures on $X$. In **[Lo$_7$]** the following was established.

5.4. THEOREM. *Suppose $\nu \in {}^*\mathscr{M}(X)$ with $|\nu|(X)$ finite. Then $\nu$ is nearstandard in the weak topology*[5] *of $\mathscr{M}(X)$, with standard part $\nu^X$ given by $\nu^X(B) = \nu_L(\mathrm{st}^{-1}(B))$ for $B \in \mathscr{B}_0$. That is, for each $f \in C(X)$*

$$\int_X f d\nu^X = {}^\circ\!\int_{{}^*X} {}^*f d\nu.$$

*If $\nu = {}^*\mu$ for some $\mu \in \mathscr{M}(X)$, then $\nu^X = \mu$.*

This result now gives a characterisation of weak convergence in $\mathscr{M}(X)$.

[5] Strictly, here and elsewhere in this subsection we should say weak* topology, but we drop the * to avoid confusion.

5.5. THEOREM. *Let $(\mu_\alpha)_{\alpha \in D}$ be a bounded net in $\mathcal{M}(X)$. Then $\mu_\alpha$ converges weakly to $\mu$ in $\mathcal{M}(X)$ if and only if for each infinite $\omega \in {}^*D$, $\mu_\omega^X = \mu$. (We say $\omega$ is* infinite *if $\omega > \alpha$ for each $\alpha \in D$.)*

To illustrate this theorem, consider measures $\mu_n$ on $[0, 1]$ where $\mu_n$ has point mass $1/n$ at each of the points $0, 1/n, \ldots, n-1/n$. We know that $\mu_n$ converges weakly to Lebesgue measure $m$ on $[0, 1]$; then Theorem 5.5 shows that $m$ is the weak standard part of $\mu_H$ for each infinite $H$, and $m = \mu_H^{[0,1]}$. This last fact was established directly in Theorem 4.1.

Loeb used the ideas in the above two theorems in a series of papers [**Lo**$_5$, **Lo**$_6$, **Lo**$_9$, **Lo**$_{10}$] on axiomatic potential theory and harmonic analysis. He obtained new results (even for the unit disc) concerning the construction of maximal representing measures for positive harmonic functions; these were constructed as the weak limits of measures using the $\mathrm{st}^{-1}$ map as outlined above.

In [**AR**] Anderson and Rashid extended these ideas to the study of convergence of nets in the dual $C(X)^d$ of $C(X)$; of course $\mathcal{M}(X)$ is a subspace of $C(X)^d$. They exploited the fact that if $\phi \in C(X)^d$ there is a *finitely* additive signed Baire measure $\nu$ on $X$ such that, for each $f \in C(X)$, $\phi(f) = \int f d\nu$. By transfer this holds for each $\phi \in {}^*(C(X)^d)$, where $\nu$ is now an internal signed *Baire measure on $^*X$. For $\nu$ with $\|\nu\|$ finite the Loeb measure $\nu_L$ on $^*X$ is used to define a signed Baire measure $\nu^X$ on $X$ by $\nu^X(B) = \nu_L(\mathrm{st}^{-1}(B))$ for $B \in \mathscr{B}_0$ just as in Theorem 5.4. Now weak convergence of a bounded net in $C(X)^d$ can be characterised as follows.

5.6. THEOREM [**AR**]. *Let $(\phi_\alpha)_{\alpha \in D}$ be a bounded net in $C(X)^d$. Under the mapping * this gives an internal net $(\phi_\alpha)_{\alpha \in {}^*D}$ in $^*(C(X)^d)$, which yields an internal net $(\nu_\alpha)_{\alpha \in {}^*D}$ of internal finitely additive signed *Baire measures as above. Then the following are equivalent:*

(a) *$\phi$ converges weakly to $\mu \in \mathcal{M}(X)$;*

(b) *for each infinite $\omega \in {}^*D$, $\nu_\omega^X = \mu$.*

This result is extended in [**AR**] to spaces that are not necessarily compact or Hausdorff, and then it is used to provide a new proof of the fact that a tight subset of $C(X)^d$ is relatively weakly compact in $\mathcal{M}_t(X)$, the set of tight measures in $\mathcal{M}(X)$.

## §6 PROBABILITY AND STOCHASTIC PROCESSES

Loeb spaces, and particularly hyperfinite Loeb spaces, have found important applications in probability and stochastic process theory, which we shall outline in the next three sections. Probabilists are interested in distributions of random variables and often do not mind too much which probability space $\mathbf{\Omega} = (\Omega, \mathscr{F}, P)$ is used to carry them and thus model the phenomena being studied[(6)]. Uniform hyperfinite Loeb spaces are universal from this point of view (Theorem 6.1). In subsections (2) and (3) we give hyperfinite Loeb space constructions of a Poisson process and Brownian motion.

[(6)] In this and subsequent sections we use $(\Omega, \mathscr{F}, P)$ to denote a probability space; we adopt probabilists' conventions such as using $Ex$ to denote $\int x dP$, and often making dependence on $\omega \in \Omega$ implicit in order to ease the notation.

§6(1) *Universality of hyperfinite Loeb spaces*

Let $M$ be a *Polish space*; that is, $M$ is a complete separable metrizable topological space. A *random variable* $x:(\Omega, \mathscr{F}, P) \to M$ is a function that is measurable with respect to the Borel $\sigma$-algebra $\mathscr{B}$ on $M$. The *distribution* of $x$ is the Borel probability $\mu$ on $M$ given by $\mu(B) = P(x^{-1}(B))$ $(= P(x \in B))$. Thus, a distribution on $M$ is just a Borel probability measure on $M$.

6.1. THEOREM. *Let $(\Omega, L(\Omega), P)$ be an infinite hyperfinite Loeb space with uniform probability measure $P = P_\Omega$. For every Polish space $M$ and distribution $\mu$ on $M$ there is a random variable $x:\Omega \to M$ whose distribution is $\mu$.*

*Proof.* This is simply a disguised version of Theorem 4.8 and the remark following it. Take $\mathbf{X} = (M, \mathscr{B}, \mu)$ and $Z = \Omega$ in that theorem; then $\mu(B) = P(\pi^{-1}(B))$ for all Borel $B \subseteq M$. So $x = \pi$ is the required random variable.

A much deeper universality theorem (for stochastic processes and adapted spaces) has been proved by Keisler and Hoover; this is described in §8.

§6(2) *Construction of a Poisson process*

In his fundamental paper [Lo$_4$] (see also [Lo$_8$]) Loeb gave an intuitive construction of a Poisson process on a uniform hyperfinite Loeb space. The idea is to define an internal discrete process on a discrete time line $T = \{0, \Delta t, 2\Delta t, \ldots\}$ where $\Delta t \approx 0$, and from it construct a standard process by taking standard parts. Loeb's internal process is obtained by considering the random arrangement of an infinite number of balls in an infinite number of intervals of infinitesimal length.

Here we will describe an alternative construction of a Poisson process (due as far as we know to Keisler). The underlying space is still hyperfinite but the Loeb probability is not the uniform one. Most of the calculations are the same as in Loeb's construction. The idea is to use the existence of infinitesimals to make precise the following intuitive description of a Poisson process $n$ with rate $\lambda$; this is a random function $n:\mathbb{R}^+ \to \mathbb{N}$ that increases by 1 in each small time interval of length $\delta t$ with probability $\lambda\delta t$, and does so independently of the past; so really $n$ is a measurable function $n:\mathbb{R}^+ \times \Omega \to \mathbb{N}$ for some probability space $\mathbf{\Omega}$.

Fix an infinite $H \in {}^*\mathbb{N}$ and let $\Delta t = H^{-1}$. Let $T = \{0, \Delta t, \ldots, H^2\Delta t\}$ and let $\Omega = \{0,1\}^{T\setminus\{0\}}$ (recall that this means all *internal* sequences of 0's and 1's indexed by $t \in T\setminus\{0\}$). Clearly $\Omega$ is hyperfinite, with $|\Omega| = 2^{H^2}$; the set $\Omega$ is an internal cartesian product. An internal measure $\nu_0$ is defined on $\{0,1\}$ by

$$\nu_0(\{1\}) = \lambda\Delta t\,; \quad \nu_0(\{0\}) = 1-\lambda\Delta t\,.$$

Define $\nu$ to be the product measure on $\Omega$ induced by $\nu_0$. Specifically, for $\omega \in \Omega$ we have

$$\Delta\nu(\omega) = p^{|\omega^{-1}(\{1\})|}(1-p)^{|\omega^{-1}(\{0\})|}$$

where $p = \lambda\Delta t$.

Now define an internal process $N: T\times\Omega \to {}^*\mathbb{N}$ by

$$N(0,\omega) = 0\,, \quad N(\mathrm{t}+\Delta t, \omega) = N(\mathrm{t},\omega)+\omega(\mathrm{t}+\Delta t)\,.$$

Thus, suppressing $\omega$ and writing $\Delta N(\mathrm{t}) = N(\mathrm{t}+\Delta t)-N(\mathrm{t})$, we see that

$$\Delta N(\mathrm{t}) = \begin{cases} 1 \text{ with probability } p = \lambda\Delta t \\ 0 \text{ with probability } 1-p, \end{cases}$$

and $\Delta N(\mathrm{t})$ is independent of $N(\mathrm{s})$ for $\mathrm{s} \leqslant \mathrm{t}$. In other words, $N$ is an internal Bernoulli process.

From the internal space $(\Omega, {}^*\mathscr{P}(\Omega), \nu)$ we obtain the standard Loeb space $(\Omega, \mathscr{F}, P)$ where $\mathscr{F} = L({}^*\mathscr{P}(\Omega))$ and $P = \nu_L$. From $N$ define a standard function $n : \mathbb{R}^+ \times \Omega \to \bar{\mathbb{N}}$ by

$$n(s, \omega) = \max_{s \approx \mathrm{t}} {}^\circ N(\mathrm{t}, \omega).$$

6.2. Theorem. *The function $n$ is a measurable Poisson process with rate $\lambda$. That is, for $k \in \mathbb{N}$ and $s, u \in \mathbb{R}^+$*

(a) $P(n_{s+u}-n_s = k) = (\lambda u)^k e^{-\lambda u}/k!$ ;

(b) *$n$ has independent increments (that is, if $0 = t_0 < t_1 < \ldots t_n$ then the random variables $n(t_{i+1}, \cdot)-n(t_i, \cdot)$ are all independent);*

(c) *$n(\cdot, \omega)$ is right continuous for every $\omega$.*

*Proof.* Fix $k \in \mathbb{N}$ and finite $\mathrm{t} \in T$; let $\mathrm{t} = K\Delta t$. By transfer of finite calculations (and suppressing $\omega$) we have

$$\begin{aligned} \nu(N(\mathrm{t}) = k) &= p^k(1-p)^{K-k}{}_KC_k \quad (\text{where } p = \lambda\Delta t) \\ &= \left(\frac{\lambda\mathrm{t}}{K}\right)^k \left(1-\frac{\lambda\mathrm{t}}{K}\right)^K (1-p)^{-k} \frac{K!}{k!(K-k)!} \quad \left(\text{since } p = \frac{\lambda\mathrm{t}}{K}\right) \\ &= \frac{(\lambda\mathrm{t})^k}{k!}\left(1-\frac{\lambda\mathrm{t}}{K}\right)^K (1-p)^{-k} \frac{K!}{K^k(K-k)!} \\ &\approx \frac{(\lambda\mathrm{t})^k}{k!} e^{-\lambda\mathrm{t}} \end{aligned}$$

since, for $K$ infinite, $\left(1-\frac{\lambda\mathrm{t}}{K}\right)^K \approx e^{-\lambda\mathrm{t}}$ and $(1-p)^{-k} \approx 1 \approx \frac{K!}{K^k(K-k)!}$. (If $K$ is finite then $\mathrm{t} \approx 0$, and note that $p^k \approx 0$.) So $P(N(\mathrm{t}) = k) = \frac{(\lambda^\circ\mathrm{t})^k}{k!} e^{-\lambda^\circ\mathrm{t}}$, and thus $N(\mathrm{t})$ is finite almost surely. A similar calculation shows that for finite $\mathrm{t}, \mathrm{u}$,

$$\nu(N(\mathrm{t}+\mathrm{u})-N(\mathrm{t}) = k) \approx \frac{(\lambda\mathrm{u})^k}{k!} e^{-\lambda\mathrm{u}};$$

this implies that for fixed $\mathrm{t}$, almost surely we have $N(\mathrm{t}) = N(\mathrm{s})$, for all $\mathrm{s} \approx \mathrm{t}$. Thus

$$P(n_{s+u}-n_s = k) = (\lambda u)^k e^{-\lambda u}/k! \quad (s, u \in \mathbb{R}^+).$$

The independence of the increments of $n$ follows easily from the corresponding fact for $N$. For right continuity of $n(\cdot, \omega)$, fix $\omega \in \Omega$ and suppress it. Since $n$ is increasing we need only consider $t$ with $n_t = k \in \mathbb{N}$. Then there is $\mathrm{t} \approx t$ with $N(\mathrm{t}) = k$ and $N(\mathrm{u}) = k$ for all $\mathrm{u} \geqslant \mathrm{t}$, $\mathrm{u} \approx \mathrm{t}$. By infinitesimal overflow there is real $\delta > 0$ with $N(\mathrm{u}) = k$ for all $\mathrm{u} \in [\mathrm{t}, \mathrm{t}+\delta)$, so $n_u = k$ for all $u \in [t, t+\delta)$.

Finally note that $n(t, \cdot)$ is measurable for each fixed $t$ (because it is lifted by $N(\mathrm{t}, \cdot)$ if $\mathrm{t} \approx t$); then using right continuity we see that $n$ is jointly measurable.

The above intuitive construction of a Poisson process allows easy verification of many of its properties, by showing that the corresponding property is evident for $N$ using simple hyperfinite combinatorics.

## §6(3) *Anderson's Brownian motion*

In the important paper [**An**$_1$] Anderson constructed Brownian motion on a hyperfinite Loeb space using the idea at the heart of the above construction of a Poisson process: define a suitable discrete internal process $B(\mathrm{t}, \omega)$ on the hyperfinite time line $T$ and then Brownian motion is obtained by taking standard parts.

Recall that a Brownian motion (in one dimension) is a measurable stochastic process $b : \mathbb{R}^+ \times \Omega \to \overline{\mathbb{R}}$ defined on some probability space $\mathbf{\Omega} = (\Omega, \mathscr{F}, P)$, such that $b$ models the behaviour of a particle whose random motion has the following characteristics.

(i) $b(0, \omega) = 0$;

(ii) for fixed $t$, $b(t, \cdot)$ is distributed normally with mean 0, variance $t$;

(iii) $b$ has independent increments;

(iv) $b(\cdot, \omega)$ is continuous for almost all $\omega$.

Anderson's construction is as follows. With $H \in {}^*\mathbb{N} \setminus \mathbb{N}$ and $T = \{0, \Delta t, \dots, H\}$ fixed as before, let $\Omega = \{-1, 1\}^{T \setminus \{0\}}$, so $\Omega$ is hyperfinite. Let $P = P_\Omega$, the uniform Loeb probability on $\Omega$. Define an internal function $B : T \times \Omega \to {}^*\mathbb{R}$ by

$$B(0, \omega) = 0\,, \quad B(\mathrm{t}+\Delta t, \omega) = B(\mathrm{t}, \omega) + \Delta t^{1/2}\, \omega(\mathrm{t}+\Delta t)\,;$$

$B$ is an internal *random walk*. Let $\Delta B(\mathrm{t}, \omega) = B(\mathrm{t}+\Delta t, \omega) - B(\mathrm{t}, \omega)$; then $\Delta B(\mathrm{t}, \omega)$ has variance $\Delta t$, which is closely connected with the requirement (ii) above. By filling in linearly $B$ can be extended to a function defined for all $t \in [0, H]$. Now we can state Anderson's result.

6.3. THEOREM [**An**$_1$]. *The standard process* $b : \mathbb{R}^+ \times \Omega \to \overline{\mathbb{R}}$ *defined by* $b(t, \omega) = {}^\circ B(t, \omega)$ *is a Brownian motion.*

*Proof.* We sketch the main points. First, use the De Moivre–Laplace central limit theorem together with Theorem 1.7(b) to see that for $\mathrm{t} \in T$ and real $\alpha$

$$\bar{P}(B(\mathrm{t}) \leqslant \alpha) \approx {}^*\Psi(\alpha/\mathrm{t}^{1/2}) \qquad (6.4)$$

where $\Psi(\alpha) = (2\pi)^{-1/2} \int_{-\infty}^{\alpha} \exp(-\frac{1}{2}x^2)dx$ is the normal distribution $N(0, 1)$. Since $\Psi$ is continuous, we see that $b(t, \cdot)$ has the distribution $N(0, t)$.

Clearly $B$ has independent increments (restricting to $\mathrm{t} \in T$) and this establishes the analogous property for $b$.

To verify that $b(\cdot, \omega)$ is continuous for almost all $\omega$, Anderson showed that for any fixed finite $\mathrm{t} \in T$, $B(\cdot, \omega)$ is almost surely $\mathscr{S}$-continuous on $[0, \mathrm{t}]$. Taking $\mathrm{t} = 1$, for example, $B(\cdot, \omega)$ is $\mathscr{S}$-continuous on $[0, 1]$ provided that $\omega \notin \bigcup_{m \in \mathbb{N}} \bigcap_{n \in \mathbb{N}} \Omega_{mn}$ where

$$\Omega_{mn} = \bigcup_{0 \leqslant i < n} \left\{\omega : \max_{\mathrm{t} \in [i/n, (i+1)/n]} B(\mathrm{t}, \omega) - \min_{\mathrm{t} \in [i/n, (i+1)/n]} B(\mathrm{t}, \omega) > 1/m\right\}.$$

Now using the symmetry properties of the random walk $B$ calculate as follows, using simple counting arguments:

$$\begin{aligned}\bar{P}(\Omega_{mn}) &\leqslant n\bar{P}\left(\max_{\mathrm{t} \in [0, 2/n]} B(\mathrm{t}) - \min_{\mathrm{t} \in [0, 2/n]} B(\mathrm{t}) > 1/m\right) \\ &\leqslant n\bar{P}\left(\max_{\mathrm{t} \in [0, 2/n]} |B(\mathrm{t})| > 1/2m\right) \\ &\leqslant 2n\bar{P}\left(\max_{\mathrm{t} \in [0, 2/n]} B(\mathrm{t}) > 1/2m\right) \\ &\leqslant 4n\bar{P}\big(B(2/n) > 1/2m\big)\end{aligned}$$

(since at least half of those $\omega$ with $\max_{\mathrm{t} \in [0, 2/n]} B(\mathrm{t}) > 1/2m$ have $B(2/n) > 1/2m$ also)

$$\approx 4n\big(1 - \Psi(n^{1/2}/2^{3/2}m)\big) \quad \text{by (6.4).}$$

From this $P(\Omega_{mn}) \to 0$ as $n \to \infty$, so $B(\cdot, \omega)$ is indeed $\mathscr{S}$-continuous on a set of full measure. Theorem 1.19 is now invoked to show that $b(\cdot, \omega)$ is continuous, almost surely.

Wiener measure (this time on $C(\mathbb{R})$) is derived from a Brownian motion $b$ by defining

$$\mu_w(A) = P(\{\omega : b(\cdot, \omega) \in A\})$$

for Borel $A \subseteq C(\mathbb{R})$. In [**An**$_1$] Anderson used his construction to prove Donsker's theorem: Wiener measure is the weak limit of the measures induced on $C(\mathbb{R}^+)$ by standard random walks $B^{(n)}$ taking steps $\pm n^{-1/2}$ in time steps of $1/n$. The ideas involved in the proof are similar to those described in §5(2).

The paper [**An**$_1$] also gives a very pleasant and intuitive development of the Itô integral, which has had important consequences; these are discussed in §7.

Anderson's Brownian motion was used by Perkins [**Pe**$_1$] to settle some open questions about Brownian local time. Briefly, for a given Brownian motion $b(t, \omega)$, the *local time* $l(t, x, \omega)$ has the property that (with $\omega$ suppressed)

$$l(t,x) = \frac{d}{dx}\int_0^t \chi_{\{b_s \leqslant x\}}\,ds \quad (t \geqslant 0, x \in \mathbb{R}),$$

(where $\chi$ is the characteristic or indicator function). That is, for Borel $A \subseteq \mathbb{R}$, we have

$$\int_A l(t,x)dx = m(\{s : s \leqslant t \,\&\, b_s \in A\})$$

(where $m$ is Lebesgue measure). Trotter established that there *is* such a function $l(t, x, \omega)$ (and that it is jointly continuous in $(t, x)$). Perkins' nonstandard construction of $l$ for Anderson's Brownian motion $b$ is as follows. The natural internal analogue of local time for the internal process $B$ is the function $L(\mathrm{t}, \mathrm{x})$ defined by

$$L(\mathrm{t}, \mathrm{x})\Delta x = \sum_{\mathrm{s}<\mathrm{t}} \chi_{\{B(\mathrm{s}) = \mathrm{x}\}}\Delta t$$

where $\mathrm{x} = k\Delta t^{1/2}$ for some $k \in {}^*\mathbb{Z}$, and $\Delta x = \Delta t^{1/2}$. This is in recognition of the fact that $B(\mathrm{t}) \in \Delta t^{1/2}\,{}^*\mathbb{Z}$, for all $\mathrm{t}$. Notice that $\omega$ occurs implicitly in the definition of $L$.

6.5. THEOREM [$\mathbf{Pe_1}$]. *For almost all $\omega$, $L(\cdot,\cdot,\omega)$ is $\mathscr{S}$-continuous and $l = {}^\circ L$ is Brownian local time.*

Many properties of local time can be obtained easily from this construction; for example, it is routine to show by induction along $T$ that

$$|B(\mathrm{t}) - \mathrm{x}| = \sum_{0 \leqslant \mathrm{s} < \mathrm{t}} \operatorname{sgn}(B(\mathrm{s}) - \mathrm{x})\Delta B(\mathrm{s}) + L(\mathrm{t}, \mathrm{x}).$$

This suggests the following formula for $l(t, x)$:

$$l(t,x) = |b(t) - x| - \int_0^t \operatorname{sgn}(b_s - x)db_s, \tag{6.6}$$

provided that some sense can be made of the final expression. The Itô integral is the way to do so, and Anderson's construction of it (see §7) shows that the above is a *proof* of Tanaka's formula (6.6).

Perkins exploited these ideas further in [$\mathbf{Pe_1}$] to obtain new results, including a global intrinsic characterisation of local time. In [$\mathbf{Pe_2}$] he showed that $l(t, x)$ is the Hausdorff $\phi$ measure of the set

$$\{s : s \leqslant t \,\&\, b_s = x\}, \text{ where } \phi(t) = (2t|\log|\log t^{-1}||)^{1/2};$$

for other applications see the papers [$\mathbf{Pe_3}$, $\mathbf{GP}$]. An informal discussion of Perkins' approach to local time is included in the survey paper [$\mathbf{Pe_5}$].

### §6(4) *White noise and Lévy Brownian motion*

Anderson's construction of Brownian motion was generalised by Stoll [St] to give an intuitive construction of white noise on a Radon measure space $\mathbf{X} = (X, \mathcal{B}, \mu)$. A white noise on $\mathbf{X}$ (with respect to a probability space $\mathbf{\Omega}$) is a random function $\lambda : \mathcal{B} \times \Omega \to \mathbb{R}$ satisfying

(a) $\lambda(B, \cdot)$ is normal with mean 0, variance $\mu(B)$, for each Borel $B$;

(b) if $B_1 \cap B_2 = \varnothing$ then $\lambda(B_1, \cdot)$ and $\lambda(B_2, \cdot)$ are independent, and $\lambda(B_1 \cup B_2, \cdot) = \lambda(B_1, \cdot) + \lambda(B_2, \cdot)$ a.s.

For $X = [0, 1]$ with Lebesgue measure, white noise is obtained from a Brownian motion $b_t(\omega)$ by defining $\lambda(C, \omega) = \int \chi_C db(\omega)$ using the stochastic integral to be described in §7. For an interval $[s, t]$ this amounts to defining $\lambda([s, t], \omega) = b_t(\omega) - b_s(\omega)$. Notice that $b$ can be recovered from $\lambda$.

Stoll's construction of a white noise on a Radon space $\mathbf{X}$ is as follows. Take the hyperfinite Loeb space representation $L(Y, {}^*\mathcal{P}(Y), \nu)$ of $\mathbf{X}$ given by Theorem 4.7. Let $\Omega = \{-1, 1\}^Y$ and define an internal function $\Lambda : {}^*\mathcal{P}(Y) \times \Omega \to {}^*\mathbb{R}$ by

$$\Lambda(A, \omega) = \sum_{y \in A} \omega(y)\sqrt{\Delta\nu(y)},$$

by analogy with Anderson's internal random walk. Clearly $\Lambda(A, \cdot)$ has $\bar{P}_\Omega$ mean zero and variance $\nu(A)$. White noise $\lambda$ on $\mathbf{X}$ is now obtained by defining $\lambda(B, \omega) = {}^\circ\Lambda(A, \omega)$ for Borel $B$, where $A \subseteq Y$ is internal with $\nu_L(A \Delta \mathrm{st}^{-1}(B)) = 0$.

Stoll used his construction of white noise to give an intuitive construction of Lévy Brownian motion on $\mathbb{R}^d$; this is a generalisation of Brownian motion $(b_t)_{t \in \mathbb{R}^+}$ in which the time set $\mathbb{R}^+$ is replaced by $\mathbb{R}^d$. A Lévy Brownian motion $(L_x)_{x \in \mathbb{R}^d}$ is a continuous process, which induces a measure on $C(\mathbb{R}^d)$ (just as Brownian motion induces Wiener measure on $C(\mathbb{R}^+)$). An important application of Stoll's construction was to obtain a Donsker-type invariance principle for the induced measure on $C(\mathbb{R}^d)$. Stoll's work is described in detail in [AFHL].

## §7 Stochastic Analysis

### §7(1) *Itô integration*

Suppose that $b$ is a Brownian motion carried by a probability space $\mathbf{\Omega} = (\Omega, \mathcal{F}, P)$: Itô showed how to make sense of an integral

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

for certain integrands $f$, where $db$ is in some sense the differential of $b$. The integral is itself a stochastic process with $I(\cdot, \omega)$ continuous a.s. We saw in §6(3) how such an integral could be useful in connection with local time; it is used extensively in modelling stochastic processes that are more general than Brownian motion.

The problem when trying to define $I(t, \omega)$ is that for almost all $\omega$, $b(\cdot, \omega)$ has unbounded variation, so no pathwise definition seems to be possible. Itô showed how $I(t, \omega)$ can be uniquely defined as a continuous process by a limiting procedure, using simple functions to approximate the integrand $f$.

The class of allowed integrands is described as follows. We must assume that $\Omega$ is equipped with additional structure, namely a collection $(\mathscr{F}_t)_{t \in \mathbb{R}^+}$ of $\sigma$-subalgebras of $\mathscr{F}$ such that

(i) $\mathscr{F}_s \subseteq \mathscr{F}_t$ if $s \leqslant t$,

(ii) $\mathscr{F}_s = \bigcap_{t > s} \mathscr{F}_t$,

(iii) $\mathscr{F}_0 \supseteq \mathscr{N}$, where $\mathscr{N}$ = the $P$-null sets in $\mathscr{F}$.

The space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \in \mathbb{R}}, P)$ is called an *adapted* (or *filtered*) probability space; $(\mathscr{F}_t)_{t \in \mathbb{R}^+}$ is a *right-continuous filtration*. Each $\sigma$-algebra $\mathscr{F}_t$ represents all that has happened up to and including time $t$. We must also assume that $b$ is an $\mathscr{F}_t$-Brownian motion; that is, $b(t, \cdot)$ is $\mathscr{F}_t$-measurable for each $t$, and the increment $b(s, \cdot) - b(t, \cdot)$ is independent of $\mathscr{F}_t$ when $s > t$.

The Itô integral above was defined for integrands $f(t, \omega)$ that are *$\mathscr{F}_t$-adapted* (that is, $f$ is measurable and $f(t, \cdot)$ is $\mathscr{F}_t$-measurable for each $t$) with $\int_0^t f^2(s, \cdot) ds < \infty$ a.s. for each $t$. The resulting integral $I(t, \omega)$ is itself adapted.

Anderson [**An**$_1$] showed how the Itô integral *can* be defined pathwise, for his Brownian motion; that is, $I(t, \omega)$ is defined separately for each $\omega \in \Omega$, using an internal Stieltjes integral (which is actually a hyperfinite sum); we will sketch his construction.

Let $(\Omega, \mathscr{F}, P)$ be the Loeb space described in §6(3) carrying Anderson's Brownian motion $b = {}^\circ B$; so $\Omega = \{-1, 1\}^{T \setminus \{0\}}$ is hyperfinite. There is a natural filtration $(\mathscr{F}_t)_{t \in \mathbb{R}^+}$ given as follows. The idea is to think of each $\omega \in \Omega$ as encoding a possible history of events for all time; then $\mathscr{F}_t$ will consist roughly of sets (events) $A$ such that whether or not $A$ occurs (i.e. whether or not $\omega \in A$) depends only on $\omega$ up to and including time $t$. In detail, fix $t \in \mathbb{R}^+$ and define $\sim_t$ on $\Omega$ by

$$\omega \sim_t \omega' \Leftrightarrow \omega \restriction \mathrm{s} = \omega' \restriction \mathrm{s} \text{ for all } \mathrm{s} \approx t$$

(here $\omega \restriction \mathrm{s}$ means $\omega \restriction \{\mathrm{u} : \mathrm{u} \leqslant \mathrm{s}\}$). Now let

$$\mathscr{F}_t^{(0)} = \{A : A \in \mathscr{F} \text{ and } A \text{ is closed under } \sim_t\}$$

and let $\mathscr{F}_t = \mathscr{F}_t^{(0)} \vee \mathscr{N}$. It is not hard to see that $(\mathscr{F}_t)_{t \in \mathbb{R}^+}$ is right continuous.

There is a natural *internal* filtration $\mathscr{A}_\mathrm{t}$ ($\mathrm{t} \in T$) given by

$$\mathscr{A}_\mathrm{t} = \{A : A \subseteq \Omega, \text{ internal, "}\omega \in A\text{" depends only on } \omega \restriction \mathrm{t}\}$$

and it is not hard to see that $\mathscr{F}_t$ can alternatively be described by

$$\mathscr{F}_t = \left(\bigcap_{t < {}^\circ\mathrm{s}} \sigma(\mathscr{A}_\mathrm{s})\right) \vee \mathscr{N} = \left(\bigcup_{t \approx \mathrm{s}} \sigma(\mathscr{A}_\mathrm{s} w)\right) \vee \mathscr{N}.$$

The resulting adapted space $\mathbf{\Omega} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}^+}, P)$ is an *adapted Loeb space*.

There is an internal counterpart to adapted functions.

7.1. DEFINITION. Let $F: T\times\Omega \to {}^*\mathbb{R}$ be internal; $F$ is *nonanticipating* if $F(\mathrm{t}, \cdot)$ is $\mathscr{A}_{\mathrm{t}}$-measurable for each t; that is, $F(\mathrm{t}, \omega)$ depends only on $\omega \upharpoonright \mathrm{t}$.

We have the following extension of the lifting theorems of §§3, 4.

7.2. THEOREM (Anderson). *Let $f: \mathbb{R}^+ \times \Omega \to \overline{\mathbb{R}}$ be a measurable function (with respect to the $\sigma$-algebra Lebesgue $\times\, \mathscr{F}$).*

(a) *There is an internal function $F: T\times\Omega \to {}^*\mathbb{R}$ such that ${}^\circ F(\mathrm{t}, \omega) = f({}^\circ\mathrm{t}, \omega)$, for almost all*[7] $(\mathrm{t}, \omega) \in \mathrm{ns}(T)\times\Omega$. *Such an $F$ is called a* lifting *of $f$.*

(b) *If $f$ is $\mathscr{F}_t$-adapted, then $f$ has a nonanticipating lifting.*

(c) *If $f$ is bounded (or integrable) there is a lifting $F$ that is bounded (or $\mathscr{S}$-integrable).*

(For a proof of this result see [**An**$_1$] or [**Ke**$_3$].)

Suppose now that $f$ is adapted and bounded with bounded nonanticipating lifting $F$. We can define an internal function $\bar{I}: T\times\Omega \to {}^*\mathbb{R}$ by

$$\bar{I}(\mathrm{t}, \omega) = \sum_{\mathrm{u}<\mathrm{t}} F(\mathrm{u}, \omega)\Delta B(\mathrm{u}, \omega).$$

(Recall that $\Delta B(\mathrm{t}, \omega) = B(\mathrm{t}+\Delta t, \omega) - B(\mathrm{t}, \omega)$.)

We can obtain a standard process from $\bar{I}$ once we have the following important result.

7.3. $\mathscr{S}$-CONTINUITY THEOREM (Keisler [**Ke**$_3$]). *Let $F$ be any internal bounded nonanticipating function; then the function $\bar{I}(\cdot, \omega)$ defined above is $\mathscr{S}$-continuous (for finite* t) *for almost all $\omega$.*

This result was established in [**An**$_1$] for the special case that $F$ is a lifting of a standard function $f$.

7.4. THEOREM [**An**$_1$]. *Let $F$ be a bounded nonanticipating lifting of a bounded adapted function $f$; then for almost all $\omega$*

$$\int_0^t f(s, \omega)\,db(s, \omega) = {}^\circ\!\left(\sum_{\mathrm{u}<t} F(\mathrm{u}, \omega)B(\mathrm{u}, \omega)\right) \quad \textit{for all } t \in \mathbb{R}^+,$$

*where the integral on the left is Itô's.*

This could be taken as a definition of the Itô integral (on this space) for bounded $f$, after showing that the expression on the right is independent of the choice of lifting $F$.

[7] Here we use the $\sigma$-finite measure induced on $\mathrm{ns}(T) = \bigcup_{n\in\mathbb{N}} T\cap[n, n+1)$ by the uniform Loeb measure on each $T\cap[n, n+1)$.

For unbounded adapted $f$ with $\int_0^t f^2(s,\cdot)ds < \infty$ a.s. for each $t$, the integral $\int f db$ is obtained by taking a nonanticipating lifting $F$ such that $F^2(\cdot,\omega)$ is $\mathscr{S}$-integrable on $[0,n]$ for all $n$, for almost all $\omega$. See [$\mathbf{An}_1$] and [$\mathbf{C}_1$] for details.

It is easy to see by internal induction along $T$ that for $\mathrm{t} \in T$

$$\bar{E}(\bar{I}^2(\mathrm{t},\omega)) = \bar{E}\left(\sum_{\mathrm{u}<\mathrm{t}} F^2(\mathrm{u},\omega)\Delta t\right),$$

since $\Delta B^2 = \Delta t$. (Here $\bar{E}$ denotes internal expectation with respect to $\bar{P}$.) This gives an easy demonstration of the fundamental isometry property of the Itô integral, namely

$$E\left(\left(\int_0^t f_s db_s\right)^2\right) = E\left(\int_0^t f_s^2 ds\right).$$

A further important property of the Itô integral is also easily derived; this is the stochastic chain rule.

7.5. Itô's Lemma (Simple version). *Let $\phi \in C^2(\mathbb{R})$ and let $x_t = \int_0^t f_s db_s$ (with $\omega$ implicit), where $f$ is $\mathscr{F}_t$-adapted. Then*

$$\phi(x_t) - \phi(x_0) = \int_0^t \phi'(x_s) f_s db_s + \tfrac{1}{2}\int_0^t \phi''(x_s) f_s^2 ds.$$

This is usually abbreviated as

$$d\phi(x_t) = \phi'(x_t)dx_t + \tfrac{1}{2}\phi''(x_t)(dx_t)^2$$

using the convention

$$dx_t = f_t db_t \quad \text{and} \quad (dx_t)^2 = f_t^2(db_t)^2 = f_t^2 dt.$$

*Proof* (Sketch). Standardly the stochastic analyst's rule of thumb "$db^2 = dt$" is only a guide; however, internally, $\Delta B^2$ *is* $\Delta t$, and this is the key to Itô's lemma. Take a suitable lifting $F$ of $f$ and then $x$ is lifted by $X$ where

$$X(\mathrm{t},\omega) = \sum_{\mathrm{u}<\mathrm{t}} F(\mathrm{u},\omega)\Delta B(\mathrm{u},\omega).$$

Fix $\omega$ such that $X$ is $\mathscr{S}$-continuous. By Taylor's theorem (transferred) we have (suppressing $\omega$)

$$\phi(X(\mathrm{t}+\Delta t)) - \phi(X(\mathrm{t})) = \phi'(X(\mathrm{t}))\Delta X(\mathrm{t}) + \tfrac{1}{2}\phi''(X(\mathrm{t})+\varepsilon(\mathrm{t}))\Delta X(\mathrm{t})^2$$

where $\Delta X(\mathrm{t}) = F(\mathrm{t})\Delta B(\mathrm{t})$ and $|\varepsilon(\mathrm{t})| < |\Delta X(\mathrm{t})| \approx 0$. Since $\phi \in C^2(\mathbb{R})$, we have $\phi(X(\mathrm{t})) \approx \phi(x({}^\circ\mathrm{t}))$, $\phi'(X(\mathrm{t})) \approx \phi'(x({}^\circ\mathrm{t}))$ and $\phi''(X(\mathrm{t})+\varepsilon(\mathrm{t})) \approx \phi''(x({}^\circ\mathrm{t}))$; thus

$$\phi(x_t)-\phi(x_0) = {}^\circ\left(\sum_{\mathrm{u}<t} \phi'(X(\mathrm{u}))\Delta X(\mathrm{u})+\tfrac{1}{2}\sum_{\mathrm{u}<t} \phi''(X(\mathrm{u})+\varepsilon(\mathrm{u}))\Delta X(\mathrm{u})^2\right)$$

$$= \int_0^t \phi'(x_s)f_s db_s+\tfrac{1}{2}\int_0^t \phi''(x_s)f_s^2 ds$$

by the nonstandard construction of the Itô and Lebesgue integrals, since $\Delta X(\mathrm{u}) = F(\mathrm{u})\Delta B(\mathrm{u})$ and so $\Delta X(\mathrm{u})^2 = F^2(\mathrm{u})\Delta B(\mathrm{u})^2 = F^2(\mathrm{u})\Delta t$.

### §7(2) *Stochastic differential equations*

Random processes can often be modelled as solutions to stochastic differential equations (SDE's) taking the general form,

$$(7.6)\qquad \begin{cases} dx_t = f(t, x_t)dt+g(t, x_t)db_t & (t \in \mathbb{R}^+) \\ x_0 = 0 \end{cases}$$

where $b$ is a Brownian motion. A solution is a continuous adapted stochastic process $x : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ such that for all $t$

$$x_t = \int_0^t f(s, x_s)ds+\int_0^t g(s, x_s)db_s;$$

thus an SDE is really an *integral* equation. Here and in (7.6) there is implicit dependence on $\omega$ throughout; the coefficients $f$ and $g$ could themselves also be dependent on $\omega$.

There are various well established techniques for solving SDE's, which depend on the nature of $f$ and $g$. In some cases the existence or otherwise of a solution depends on the particular underlying adapted space $\mathbf{\Omega}$ and the particular Brownian motion $b$ on it; to obtain a solution it may be necessary to move to another space $\mathbf{\Omega}'$ and Brownian motion $b'$; such a solution is called a *weak* solution.

In the important memoir [**Ke**$_3$] Keisler initiated a new approach to SDE's using Anderson's construction of the Itô integral. Taking $\mathbf{\Omega}$ to be the adapted Loeb space described in the previous subsection, Keisler showed that under conditions similar to or weaker than those needed to give weak solutions in general, an SDE on $\mathbf{\Omega}$ has a *strong* solution; that is, a solution on $\mathbf{\Omega}$ with $b$ prescribed.[8] Moreover, solutions on $\mathbf{\Omega}$ are obtained with remarkable ease, as we shall see.

Keisler first showed that it is sufficient to look for solutions for Anderson's Brownian motion.

[8] In the literature a *strong* solution sometimes means a solution that is adapted to the filtration generated by $b$; to avoid confusion it has been suggested that this should be called a *strict* solution; see [**Ba**] and [**C**$_1$].

7.7. THEOREM [**Ke**$_3$]. *On the adapted Loeb space* $\mathbf{\Omega}$, *if the* SDE (7.6) *with* $f, g$ *non-random has a solution for one particular Brownian motion on* $\mathbf{\Omega}$ *then it has one for every Brownian motion on* $\mathbf{\Omega}$.

We now illustrate Keisler's method with an example of a new strong existence theorem from [**Ke**$_3$].

7.8. THEOREM. *Suppose that the functions* $f, g : \mathbb{R}^+ \times \mathbb{R} \times \Omega \to \mathbb{R}$ *are bounded, measurable and*

(i) *for each* $(t, x)$, $f(t, x, \cdot)$ *and* $g(t, x, \cdot)$ *are* $\mathscr{F}_t$*-measurable,*

(ii) *for each* $(t, \omega)$, $f(t, \cdot, \omega)$ *and* $g(t, \cdot, \omega)$ *are continuous. Then the* SDE

$$dx_t = f(t, x_t)dt + g(t, x_t)db_t, \quad x_0 = 0$$

*has a solution for any Brownian motion* $b$ *on* $\mathbf{\Omega}$.

*Proof* (slightly extended). By Theorem 7.7, we need only consider $b = {}^\circ B =$ Anderson's Brownian motion. Writing $\hat{f}(t, \omega) = f(t, \cdot, \omega)$ and $\hat{g}(t, \omega) = g(t, \cdot, \omega)$ we have $\hat{f}, \hat{g} : \mathbb{R}^+ \times \Omega \to C(\mathbb{R})$, and $\hat{f}, \hat{g}$ are adapted. The space $C(\mathbb{R})$ with the compact-open topology is separable and metrizable, so the adapted lifting theorem (7.2) generalises to give internal bounded nonanticipating functions $\hat{F}, \hat{G} : T \times \Omega \to {}^*C(\mathbb{R})$ such that $\hat{F}(\mathrm{t}, \omega) \approx \hat{f}({}^\circ\mathrm{t}, \omega)$ and $\hat{G}(\mathrm{t}, \omega) \approx \hat{g}({}^\circ\mathrm{t}, \omega)$ for almost all $(\mathrm{t}, \omega) \in \mathrm{ns}(T) \times \Omega$. Here $\approx$ refers to the compact-open topology on $C(\mathbb{R})$. Now define internal $F, G : T \times {}^*\mathbb{R} \times \Omega \to {}^*\mathbb{R}$ by

$$F(\mathrm{t}, x, \omega) = \big(\hat{F}(\mathrm{t}, \omega)\big)(x), \qquad G(\mathrm{t}, x, \omega) = \big(\hat{G}(\mathrm{t}, \omega)\big)(x).$$

Then $F, G$ are bounded, and by properties of nearstandardness in $C(\mathbb{R})$, for almost all $(\mathrm{t}, \omega) \in \mathrm{ns}(T) \times \Omega$,

$$F(\mathrm{t}, x, \omega) \approx f({}^\circ\mathrm{t}, {}^\circ x, \omega), \qquad G(\mathrm{t}, x, \omega) \approx g({}^\circ\mathrm{t}, {}^\circ x, \omega)$$

for all finite $x \in {}^*\mathbb{R}$.

The internal counterpart of the SDE is a hyperfinite difference equation whose solution $X : T \times \Omega \to {}^*\mathbb{R}$ is given by

$$X(0, \omega) = 0,$$

$$X(\mathrm{t} + \Delta t, \omega) = X(\mathrm{t}, \omega) + F\big(\mathrm{t}, X(\mathrm{t}, \omega), \omega\big)\Delta t + G\big(\mathrm{t}, X(\mathrm{t}, \omega), \omega\big)\Delta B(\mathrm{t}, \omega).$$

The $\mathscr{S}$-continuity theorem (7.3) shows that $X(\cdot, \omega)$ is $\mathscr{S}$-continuous for almost all $\omega$; so a standard process $x$ is defined for a.a. $\omega$ by

$$x({}^\circ\mathrm{t}, \omega) = {}^\circ X(\mathrm{t}, \omega) \quad (\mathrm{t} \text{ finite}).$$

The above shows that $F\big(\mathrm{t}, X(\mathrm{t}, \omega), \omega\big)$ and $G\big(\mathrm{t}, X(\mathrm{t}, \omega), \omega\big)$ lift $f\big(t, x(t, \omega), \omega\big)$ and $g\big(t, x(t, \omega)\big)$, so by the nonstandard characterisation of the Lebesgue and Itô integrals we have for a.a. $\omega$, if $t \in \mathbb{R}^+$ and $t \approx \mathrm{t}$,

$$x_t = {}^\circ X(\mathrm{t}) = {}^\circ\left(\sum_{\mathrm{u} < \mathrm{t}} F(\mathrm{u}, X(\mathrm{u}))\Delta t + \sum_{\mathrm{u} < \mathrm{t}} G(\mathrm{u}, X(\mathrm{u}))\Delta B(\mathrm{u})\right)$$

$$= \int_0^t f(s, x_s)ds + \int_0^t g(s, x_s)db_s,$$

so $x$ is a solution to the SDE.

Further strong existence theorems are proved in [**Ke**$_3$], with alternative conditions on $f, g$. For example, solutions are obtained for non-random $f, g$ without requiring continuity in $x$, provided that $g^{-1}$ is bounded. An improved proof is given in [**Li**$_3$]. The methods work in higher dimensions and more general initial conditions can be handled.

In [**C**$_1$] and [**HP**] Keisler's approach was extended to more general equations in which the coefficients $f, g$ may depend at time $t$ on the past history $(x_s)_{s \leqslant t}$ of $x$. The results of [**C**$_1$] were subsequently applied to stochastic control theory; see §9(2) below.

§7(3) *Stochastic integration*

There is a well developed standard theory in which Itô integration (where the integrator is a Brownian motion) is extended to a theory of integration against martingales and even semimartingales. A nonstandard approach to this general form of stochastic integration was developed independently by Hoover and Perkins [**HP**], and Lindstrøm [**Li**$_1$]. For continuous martingales this was first worked out by Panetta [**Pa**].

The standard theory develops a stochastic integral $I(t, \omega) = \int_0^t f(s, \omega)dm(s, \omega)$ where $m : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ is a *semimartingale* and $f : \mathbb{R}^+ \times \Omega \to \mathbb{R}$ is a process that is *predictable*; this is the appropriate generalisation of the notion of adaptedness. The semimartingale $m$ has $m(\cdot, \omega) \in D(\mathbb{R}^+)$ for almost all $\omega$, where $D(\mathbb{R}^+)$ denotes the space of real functions on $\mathbb{R}^+$ that are right continuous with left limits. The integral $I(t, \omega)$ has the same property.

The pattern of the nonstandard development of stochastic integration is similar to that given by Anderson for the Itô integral, but the details are much more complicated (as in the standard theory).

Hoover and Perkins work with a general adapted Loeb space rather than the specific adapted space described in §7(1) above; such a space is obtained by taking an arbitrary *internal* adapted space $(\Omega, \mathscr{B}, (\mathscr{B}_\mathrm{t})_{\mathrm{t} \in T}, \nu)$ where $(\mathscr{B}_\mathrm{t})_{\mathrm{t} \in T}$ is an internal filtration for the discrete time set $T$, so that $\mathscr{B}_\mathrm{t} \subseteq \mathscr{B}_\mathrm{s} \subseteq \mathscr{B}$ whenever $\mathrm{t} < \mathrm{s}$; $\Omega$ is not necessarily hyperfinite. The standard adapted Loeb space is then

$$\mathbf{\Omega} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t \in \mathbb{R}^+}, P)$$

where $P = \nu_L$, $\mathscr{F} = L(\mathscr{B})$ and, with $\mathscr{N}$ = null sets in $\mathscr{F}$,

$$\mathscr{F}_t = \left(\bigcap_{t < {}^\circ\mathrm{s}} \sigma(\mathscr{B}_{\mathrm{s}})\right) \vee \mathscr{N} = \left(\bigcup_{t \approx \mathrm{s}} \sigma(\mathscr{B}_{\mathrm{s}})\right) \vee \mathscr{N}.$$

Hoover and Perkins [**HP**] adopt the point of view that "... adapted Loeb spaces suffice for practical purposes. In particular, we believe that typical physical processes can be modelled on them directly". Evidence to substantiate such remarks is provided in the paper [**HK**] which we discuss in §8.

Lindstrøm restricts his attention to hyperfinite Loeb spaces and shows that these are sufficiently general to represent any standard adapted space in a weak sense similar to that described in Theorem 4.9. In most respects Lindstrøm's approach follows the same lines as that in [**HP**].

For the Lindstrøm–Hoover–Perkins programme it is necessary to have an internal representations of functions $f \in D(\mathbb{R}^+)$ that corresponds to the representation of continuous functions by $\mathscr{S}$-continuous functions.

7.9. Definition [**HP**]. An internal function $G : T \to {}^*\mathbb{R}$ is SDJ if

(a) $G(\mathrm{t})$ is finite, for all finite t;

(b) $G(\mathrm{t}) \approx G(0)$ if $\mathrm{t} \approx 0$;

(c) for each $t \in \mathbb{R}^+$ there is $\mathrm{t} \approx t$ such that $G(\mathrm{s}) \approx G(\mathrm{u})$ whenever $\mathrm{s}, \mathrm{u} \approx t$ and *either* $\mathrm{s}, \mathrm{u} < \mathrm{t}$ *or* $\mathrm{s}, \mathrm{u} \geqslant \mathrm{t}$.

It is easy to establish the following (cf. Theorems 1.19 and 1.6.)

7.10. Theorem [**HP**].

(a) *If $G$ is* SDJ *then the function $g$ given by*

$$g(t) = \sup_{\mathrm{s} \approx t} {}^\circ G(\mathrm{s}) \quad (t \in \mathbb{R}^+)$$

*is in $D(\mathbb{R}^+)$ and $g(t-) = \inf_{\mathrm{s} \approx t} {}^\circ G(\mathrm{s})$.*

(b) *If $g \in D(\mathbb{R}^+)$, then ${}^*g \restriction T$ is* SDJ.

The stochastic integral $\int_0^t f\,dm$, where $f$ is predictable and $m$ is a semimartingale, is obtained in [**HP**, **Li$_1$**] essentially by taking suitable liftings $F$ and $M$ of $f, m$, so that $M(\cdot, \omega)$ is SDJ for almost all $\omega$; then it is shown that for a.a. $\omega$

$$\int_0^t f\,dm = {}^\circ\sum_{\mathrm{s} < t} F(\mathrm{s})\Delta M(\mathrm{s}),$$

where $\Delta M(\mathrm{t}, \omega) = M(\mathrm{t} + \Delta t, \omega) - M(\mathrm{t}, \omega)$.

In [**HP**] Hoover and Perkins go on to apply their results to obtain new results about the solutions of SDE's with semimartingales as integrators; the essence of their method is that of Keisler [**Ke**$_3$], described in §7(2).

The survey paper [**Pe**$_5$] contains a more detailed presentation of the key ideas and results involved in the work described in this section.

In another generalisation of Anderson's Itô integration, Lindstrøm [**Li**$_4$] developed the theory of stochastic integration in hyperfinite dimensional spaces. He thus obtained an Anderson type approach to Brownian motions associated with Hilbert spaces, and the corresponding theory of Itô integration. This yields results about the solution of partial SDE's in such a context.

### §7(4) *Other topics in stochastic analysis*

For the sake of comprehensiveness we mention some further papers on topics related to those described in this section. In the paper [**Ko**$_1$] Kosciuk develops an adapted Loeb space approach to the solution of a class of martingale problems, thereby obtaining solutions to stochastic partial differential equations that model the diffusion of enzymes in the presence of DNA molecules; such ideas form the subject matter of Kosciuk's thesis [**Ko**$_2$].

In [**Pe**$_4$] Perkins used an adapted Loeb space to construct a local martingale whose absolute value has a prescribed law.

Finally, Helms and Loeb [**HL**$_3$] gave a neat nonstandard proof of the fundamental martingale convergence theorem, using pre-Loeb measure techniques.

Of course the books [**AFHL**] and [**BS**] contain a wealth of material on Loeb space techniques in stochastic analysis and, in [**AFHL**], a wide range of physically oriented applications—some of which we shall touch on in §9.

## §8 Adapted Distributions

Here we describe some of the results in a recent preprint of Hoover and Keisler [**HK**] that provides evidence for the view that it is not restrictive to use adapted Loeb spaces on which to model stochastic processes. This has been expressed informally by saying that anything that happens anywhere happens on a Loeb space. A simple formulation of this for random variables was provided by Theorem 6.1.

Suppose that $x, y$ are measurable stochastic processes[9] living on adapted spaces $\mathbf{\Omega} = (\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}^+}, P)$ and $\mathbf{\Lambda} = (\Lambda, \mathscr{G}, (\mathscr{G})_{t\in\mathbb{R}^+}, Q)$ respectively, taking values in some Polish space $M$. Then $x$ and $y$ are said to have the same *adapted distribution* if all processes obtained from $x, y$ by iteration of (i) composition with continuous functions and (ii) taking conditional expectations (with respect to $\mathscr{F}_t$ or $\mathscr{G}_t$) have the same finite dimensional distributions. This notion is written $x \equiv y$; it is made precise in Definition (8.3) below.

The main theorem of [**HK**] involves the notions of *universality* and *saturation* for adapted spaces.

[9] Throughout this section all stochastic processes are assumed to be measurable; that is, measurable in the product $\mathbb{R}^+ \times \Omega$.

8.1. DEFINITION. (a) An adapted space $\mathbf{\Omega}$ is *universal* if for every Polish space $M$ and every $M$-valued stochastic process $y$ on an adapted space $\mathbf{\Lambda}$, there is an $M$-valued process $x$ on $\mathbf{\Omega}$ such that $x \equiv y$.

(b) The space $\mathbf{\Omega}$ is *saturated* if for every Polish space $M$ and every pair of $M$-valued processes $y_1, y_2$ on an adapted space $\mathbf{\Lambda}$ and every $x_1$ on $\mathbf{\Omega}$ with $x_1 \equiv y_1$ there is a process $x_2$ on $\mathbf{\Omega}$ such that $(x_1, x_2) \equiv (y_1, y_2)$.

Clearly saturation implies universality. The next theorem is the main result of [**HK**].

8.2. THEOREM. *If $\mathbf{\Omega}$ is an adapted Loeb space that carries a Brownian motion then $\mathbf{\Omega}$ is saturated.*

Adapted Loeb spaces are the only saturated spaces known to Hoover and Keisler.

The idea of the adapted distribution of a process is made precise as follows. For a fixed Polish space $M$ first define the class $CP = \bigcup_{k \in \mathbb{N}} CP^{(k)}$ of *conditional processes* in the following way. Each $f \in CP^{(k)}$ is an operator on processes $x$ (on an adapted space $\mathbf{\Omega}$) so that $fx$ is a $k$-fold real process living on the same space; that is, $fx$ is a measurable function $fx : (\mathbb{R}^+)^k \times \Omega \to \mathbb{R}$. The class $CP$ is built up inductively, with the following clauses.

(i) If $\Phi : M^k \to \mathbb{R}$ is bounded, continuous, then $\hat{\Phi} \in CP^{(k)}$ where

$$(\hat{\Phi}x)_{t_1, t_2, \dots, t_k} = \Phi(x_{t_1}, \dots, x_{t_k}) .$$

(ii) If $\phi : \mathbb{R}^m \to \mathbb{R}$ is bounded, continuous, and $f_1, \dots, f_m \in CP^{(k)}$ then $\phi(f_1, \dots, f_m) \in CP^{(k)}$ where

$$(\phi(f_1, \dots, f_m)x)_{\mathbf{t}} = \phi((f_1 x)_{\mathbf{t}}, (f_2 x)_{\mathbf{t}}, \dots, (f_m x)_{\mathbf{t}})$$

(writing $\mathbf{t} = (t_1, \dots, t_k)$).

(iii) If $f \in CP^{(k)}$ then $E[f \mid -] \in CP^{(k+1)}$ where

$$(E[f \mid -]x)_{t, t_1, \dots, t_k} = E((fx)_{t_1, \dots, t_k} \mid \mathscr{F}_t).$$

8.3. DEFINITION. Let $x, y$ be $M$-valued processes living on adapted spaces $\mathbf{\Omega}, \mathbf{\Lambda}$; then $x$ and $y$ have the same *adapted distribution*, written $x \equiv y$, if whenever $f \in CP^{(k)}$ and $\mathbf{t} \in (\mathbb{R}^+)^k$ then $E((fx)_{\mathbf{t}}) = E((fy)_{\mathbf{t}})$.

Clearly $\equiv$ is a very strong equivalence relation; it seems to preserve most properties of interest to stochastic analysts. Thus the saturation property of adapted Loeb spaces indicates strongly that working with Loeb spaces is not a restriction.

The following application of saturation shows why solving SDE's on adapted Loeb spaces is so successful.

8.4. THEOREM [**HK**]. *Suppose that $m, \hat{m}$ are semimartingales on spaces $\mathbf{\Lambda}, \mathbf{\Omega}$ with $m \equiv \hat{m}$, and let $\phi$ be a Borel function. Suppose that $x$ is a solution on $\mathbf{\Lambda}$ to the* SDE

$$x_t = \int_0^t \phi(x_{s-})dm_s.$$

*Then if $\mathbf{\Omega}$ is saturated there is a process $\hat{x}$ on $\mathbf{\Omega}$ such that $(\hat{m}, \hat{x}) \equiv (m, x)$ and $\hat{x}$ is a solution to*

$$\hat{x}_t = \int_0^t \phi(\hat{x}_{s-})d\hat{m}_s.$$

Model theoretic aspects of adapted distributions have been developed by Keisler and studied in [**Ke$_4$**] and [**Rd$_2$**]. The question addressed by Rodenhausen can be formulated as follows (with time restricted to $[0, 1]$). Note that a process $x$ on an adapted space $\mathbf{\Lambda}$ defines a *valuation* $v_x: CP \to \mathbb{R}$ by

$$v_x(f) = \int_{0 \leqslant t \leqslant 1} (fx)_t dt.$$

Rodenhausen asks: what are necessary and sufficient conditions on an arbitrary $v: CP \to \mathbb{R}$ for it to be $v_x$ for some process $x$ on some adapted space $\mathbf{\Lambda}$? The completeness theorem of [**Rd$_2$**] establishes necessary and sufficient conditions in the form of axioms and rules; this gives a real valued logical system, whose models (or interpretations) are given by processes $x$. To establish the sufficiency of his conditions Rodenhausen constructs a process $x$ as $x = {}^\circ X$ where $X$ is an internal process on an internal adapted space, so $x$ lives on the corresponding adapted Loeb space.

## §9 FURTHER APPLICATIONS

We conclude our survey with a brief review of some diverse applications of Loeb measures not discussed in earlier sections.

### §9(1) *Self-avoiding random walks*

Random walks having paths that do not cross themselves have been proposed as models of long polymer chains. A new approach was taken by Lawler [**La**] who studied those self-avoiding random walks obtained by erasing loops in an ordinary random walk. Applying this to Anderson's internal random walk with Brownian motion as its standard part, Lawler constructed a *self-avoiding Wiener walk.* In [**La**] the following was then established.

9.1. THEOREM. *Let $S(k, \omega)$ $(k = 0, 1, \ldots)$ be a simple random walk in d-dimensions, with each step a unit vector parallel to one of the coordinate axes; let $\hat{S}(t, \omega)$ be obtained from $S$ by erasing loops and filling in linearly. Define a scaled self-avoiding process $\hat{Y}_n$ by $\hat{Y}_n(t, \omega) = dn^{-1/2}\hat{S}(nt, \omega)$. Then if $d \geqslant 5$ there is a constant $\kappa_d$ such that $\hat{Y}_n(\kappa_d t, \omega)$ converges in distribution to the Wiener distribution.*

This result should be compared with Donsker's Theorem (see §6(3)) which says that if $S(k, \omega)$ is filled in linearly and similarly scaled, then the Wiener distribution is obtained in the limit.

### §9(2) *Optimal control theory*

In deterministic control theory, the trajectory $x_t$ of an object being controlled is often described by a differential equation such as

$$dx_t = f(t, x_t, u_t)dt \quad (0 \leqslant t \leqslant 1\,; x_0 = 0), \tag{9.2}$$

where $f$ is bounded and the control function $u$ is a measurable function $u : [0, 1] \to A$, a compact subset of Euclidean space. A cost $J(u)$ is specified for each control $u$; for example a terminal cost specifies that $J(u) = c(x_1)$, where $c$ is a fixed positive continuous function. An *optimal* control is one that minimises $J(u)$.

A natural nonstandard approach to finding an optimal control is as follows. Let $(u_n)_{n \in \mathbb{N}}$ be a sequence of controls with $\inf(J(u_n)) = \inf(J(u)) = J_{\min}$, say. Take liftings $U_n : T \to {}^*A$ of $u_n$, and take a suitable lifting $F : T \times {}^*\mathbb{R} \times {}^*A \to {}^*\mathbb{R}$ of the function $f$. (Here $T$ is the discrete time set $\{0, \Delta t, \dots, 1\}$.) Define $X_n : T \to {}^*\mathbb{R}$ by

$$X_n(\mathrm{t}+\Delta t) = X_n(\mathrm{t}) + F(\mathrm{t}, X_n(\mathrm{t}), U_n(\mathrm{t}))\Delta t\,.$$

Putting $x_n = {}^\circ X_n$ it is routine to show that $x_n$ is a solution to (9.2) for control $u_n$, and that $J(u_n) \approx {}^*c(X_n(1)) = J(U_n)$ say. Using $\omega_1$-saturation, we obtain infinite $H$ such that $J(U_H) \approx J_{\min}$. Let $U = U_H$; $U$ is an *internal* optimal control; however $U$ need not correspond to any standard control $u$, since ${}^\circ U(\mathrm{t})$ may vary for $\mathrm{t} \approx t$. In $[C_2]$ it was shown that $U$ corresponds to a *generalised* or *relaxed* control $v : [0, 1] \to \mathscr{M}(A)$, where $\mathscr{M}(A)$ is the set of Borel measures on $A$. Roughly, for a Borel set $B \subseteq A$, $v_t(B)$ is the conditional probability

$$v_t(B) = P({}^\circ U(\mathrm{t}) \in B \mid \mathrm{t} \approx t)\,.$$

The idea is made precise using the uniform Loeb measure $P$ on $T$ together with regular conditional distributions (or the disintegration of measures). The sense in which $v$ corresponds to $U$ is given by the main theorem of $[C_2]$.

9.3. THEOREM. *Let $U : T \to {}^*A$ be an internal control; there is a relaxed control $v : [0, 1] \to \mathscr{M}(A)$ such that for every bounded measurable $g : [0, 1] \times A \to \mathbb{R}$, if $g(t, \cdot)$ is continuous for each $t$, and if $G$ is a bounded uniform lifting of $g$ then*

$$\int_0^t g(s, v_s)ds = {}^\circ\!\sum_{\mathrm{t} < t} G(\mathrm{t}, U(\mathrm{t}))\Delta t\,.$$

(*A* uniform *lifting $G$ has the property that* $G(\mathrm{t}, a) \approx g({}^\circ\mathrm{t}, {}^\circ a)$ *for* all $a \in {}^*A$, *for almost all* $\mathrm{t} \in T$. *To make sense of the left hand side of this equation, the domain of $g$ is extended by setting* $g(t, \mu) = \int_A g(t, a)d\mu(a)$ *for* $\mu \in \mathscr{M}(A)$.)

Returning to the controlled system under discussion, let $U = U_H$ as above and consider the standard trajectory $x = {}^\circ X_H$ (which is defined because $X_H$ is $\mathscr{S}$-continuous); applying Theorem 9.3 to $g(t, a) = f(t, x_t, a)$ and its lifting $G(\mathrm{t}, a) = F(\mathrm{t}, X_\mathrm{t}, a)$ we see that

$$\int_0^t f(s, x_s, v_s)ds = {}^\circ\sum_{\mathrm{t}<t} F(\mathrm{t}, X_\mathrm{t}, U(\mathrm{t}))\Delta t = x_t,$$

so $x$ is the solution to (9.2) for control $v$. Moreover, $J(v) \approx J(U_H)$. Thus $v$ is optimal, and in fact $v$ is optimal amongst relaxed controls too. In some cases an ordinary optimal control can be obtained from $v$.

The above approach was introduced into stochastic control theory in [$\mathbf{C}_3$, $\mathbf{C}_4$], where new results were obtained. A controlled stochastic system is modelled by an SDE such as

$$dx_t = f(t, x, u_t)dt + g(t, x)db_t \quad (0 \leqslant t \leqslant 1\,; x_0 = 0)$$

where $x_t \in \mathbb{R}^d$ and $b$ is a $d$-dimensional Brownian motion. The coefficients $f, g$ depend on the past $(x_s)_{s \leqslant t}$ of the trajectory $x$ of the object being controlled. The control $u$ at time $t$ may depend with advantage on the past $(x_s)_{s \leqslant t}$; this is a *feedback* control. In a partially observed system this dependence is through some partial observation $y_t = \phi((x_s)_{s \leqslant t})$ of the past of $x$; for example $y_t$ could be some projection of $x_t$. In [$\mathbf{C}_3$] the Loeb space approach to the solution of SDE's (§7(2)) was combined with the above method of obtaining relaxed controls, to give new existence results for optimal controls in partially observed stochastic systems; an extension is mentioned in [$\mathbf{C}_4$].

## §9(3) *Infinite particle systems*

(i) *Statistical Mechanics.* In the papers [$\mathbf{Hu}_1$, $\mathbf{Hu}_2$] Hurd introduced Loeb measures into the study of the foundations of statistical mechanics. There are various accepted standard ways to construct an infinite *configuration space*, a measurable space $(X, \mathscr{F})$ that is in some sense a limit of large finite configuration spaces $(X_i, \mathscr{F}_i)_{i\in I}$ carrying measure theoretic phenomena that are taken to model the world as it actually exists. The role of the ideal structure $(X, \mathscr{F})$ is to make limiting phenomena in the spaces $(X_i, \mathscr{F}_i)$ more amenable. Hurd's approach was to use an infinite hyperfinite space $X$ with $\mathscr{F} = \sigma(^*\mathscr{P}(X))$ as the limiting space. Limiting measure theoretic phenomena are represented by Loeb measures on $X$. He showed that properties or questions about various families of entities (e.g. entropy, pressure, etc.,) on the spaces $(X_i, \mathscr{F}_i)$ become properties or questions about a corresponding entity on $(X, \mathscr{F})$, and vice versa. In the paper [$\mathbf{Hu}_2$] Hurd investigated the time evolution of measures on such spaces.

(ii) *Spin models.* In [$\mathbf{HL}_1$] Helms and Loeb model an infinite spin system in $d$-dimensions as follows. Take an infinite hyperinteger $H$ and let $\Gamma$ be the lattice $\Gamma = \{-H, -H+1, \ldots, -1, 0, 1, \ldots, H\}^d$. Let $S = \{-1, 1\}^\Gamma$, the set of all *internal* functions $\eta : \Gamma \to \{-1, 1\}$. Each $\eta$ is a (*spin*) *configuration*; $\eta(x)$ denotes the spin $(\pm 1)$ at the site $x \in \Gamma$.

The idea is that $\eta$ evolves with time in a stochastic manner governed by a *speed function* $c(x, \eta)$: if the system has configuration $\eta$ at time t then $c(x, \eta)\Delta t$ is the probability that the spin $\eta(x)$ is reversed at time $t + \Delta t$. This defines the internal evolution of the configuration $\eta(x)$, which in turn gives an internal Markov semigroup.

Helms and Loeb showed that the internally evolving system described above can be standardised using Loeb measure to give a Markov semigroup that is a standard description of a standard stochastically evolving infinite spin system. Thus they established that infinite spin systems can be modelled by hyperfinite systems having all the formal properties of finite spin systems. This idea was pursued further in [**HL**$_2$]. In the paper [**H1**] Helms developed a hyperfinite representation for infinite spin systems with unlimited range interaction between sites; by standardising this (again using the Loeb construction) he obtained standard martingale models for such systems.

The book [**AFHL**] will contain a detailed discussion of hyperfinite lattice models, including a derivation of the global Markov property.

(iii) *Quantum field theory.* In [**F**] Fenstad describes a Loeb measure approach to Euclidean quantum field theory. The free Euclidean field is obtained from an appropriate internal random field on a hyperfinite lattice with infinitesimal spacing (cf. (ii) above where the lattices have discrete spacing). This explicit and pointwise definition is very helpful for further constructions and computations. Full details of this approach will appear in [**AFHL**].

## §9(4) *Mathematical economics*

The idea of using a hyperfinite set to model a large finite set was used in economics by Keisler in the preprint [**Ke**$_5$]. He constructed an exchange economy with a hyperfinite set of traders $A$, each acting independently. Each trader $a \in A$ trades at time t with probability $I_a \Delta t$, where $I_a$ is an *intensity function* that depends on current prices and stock held by $a$. A natural internal probability space is constructed, and the internal evolution of the economy is easily described. A standard stochastic exchange economy is obtained by taking the appropriate Loeb space. Keisler obtained in [**Ke**$_5$] results about the predictability of the movement of prices and the convergence of prices to an equilibrium.

## §9(5) *The Boltzmann equation*

In the papers [**Ar**$_1$, **Ar**$_2$, **Ar**$_3$, **Ar**$_4$] Arkeryd introduced nonstandard ideas into the study of the Boltzmann equation for the time evolution of a gas density. For the non-linear spatially dependent equation, Loeb space techniques were used in [**Ar**$_1$] to provide a new kind of solution—a Loeb-$L^1$-sub-solution. In [**Ar**$_2$] it was shown that Loeb-solutions to the time-wise approximated Boltzmann equation converge in a weak sense to the Loeb-solutions obtained in [**Ar**$_1$].

The papers [**Ar**$_3$, **Ar**$_4$] used nonstandard methods to deal with the non-linear space homogeneous Boltzmann equation. Standard existence results for finite range intermolecular forces were extended in [**Ar**$_3$] to give solutions for the case of infinite range forces. In [**Ar**$_4$] the convergence of these solutions as $t \to \infty$ was established.

Arkeryd's work will be discussed in [**AFHL**].

§9(6) *The ergodic theorem*

We mention finally a simple proof of the ergodic theorem obtained by Kamae [**Ka**] using Loeb measure. The idea is to first establish the ergodic theorem for the system $(\mathbf{Y}, \phi)$ where $\mathbf{Y} = (Y, L(Y), P_Y)$ is a uniform hyperfinite Loeb space with $Y = \{1, 2, ..., H\}$ for some infinite $H$, and $\phi(y) = y+1 \pmod H$. Then the general ergodic theorem for a system $(\mathbf{\Omega}, \psi)$ (where $\psi$ is a measure preserving transformation on a probability space $\mathbf{\Omega}$) is obtained by showing that $(\mathbf{\Omega}, \psi)$ is a factor of $(\mathbf{Y}, \phi)$.

## Appendix: Construction of ${}^*V(\mathbb{R})$

The following construction of a nonstandard universe ${}^*V(\mathbb{R})$ was given in [**Ke$_6$**]; it is essentially the same as that in [**Ke$_2$**] and [**SL**]. To simplify the exposition we begin by omitting the enlargement axiom.

THEOREM. *There is a nonstandard universe* ${}^*V(\mathbb{R}) \subseteq V({}^*\mathbb{R})$ *with mapping* ${}^* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ *satisfying Axioms* 1.9. I, II, III.

*Proof.* Let $I$ be a countable index set and let $\mathcal{U}$ be a nonprincipal ultrafilter on $I$; that is, $\mathcal{U}$ extends the filter of cofinite subsets of $I$. For $n \in \mathbb{N}$ define

$$W_n = \{a \in V(\mathbb{R})^I : a(i) \in V_n(\mathbb{R}) \text{ a.c.}\}$$

where a.c. (*almost certainly*) means that $\{i : a(i) \in V_n(\mathbb{R})\} \in \mathcal{U}$; let

$$W_\infty = \bigcup_{n \in \mathbb{N}} W_n .$$

Let $W_\infty/\mathcal{U}$ denote the quotient of $W_\infty$ by the equivalence relation $\sim$ given by $a \sim b \Leftrightarrow a(i) = b(i)$ a.c. A pseudo membership relation $e$ is defined on $W_\infty/\mathcal{U}$ by $(a/\mathcal{U})e(b/\mathcal{U}) \Leftrightarrow a(i) \in b(i)$ a.c., after checking that this is well-defined. The resulting *bounded ultrapower* $(W_\infty/\mathcal{U}, e)$ is a structure that is isomorphic to the structure $({}^*V(\mathbb{R}), \in)$ that we are aiming to construct; we shall see that $(W_n/\mathcal{U}, e)$ is isomorphic to $({}^*V_n(\mathbb{R}), \in)$.

To obtain ${}^*V(\mathbb{R})$ from $W_\infty/\mathcal{U}$ carry out the following steps.

(1) Define ${}^*\mathbb{R} = W_0/\mathcal{U}$, and identify $\mathbb{R}$ with the set $\{\mathbf{r}/\mathcal{U} : r \in \mathbb{R}\}$, where $\mathbf{c}$ denotes the function that is constant with value $c$, for any $c \in V(\mathbb{R})$.

(2) Define a mapping $\hat{}: W_\infty \to V({}^*\mathbb{R})$ inductively as follows:

(i) $\hat{a} = a/\mathcal{U}$ if $a \in W_0$;

(ii) assuming that $\hat{}$ is defined on $W_n$ define

$$\begin{aligned} \hat{a} &= \{\hat{b} : b(i) \in a(i) \text{ a.c.}\} \\ &= \{\hat{b} : (b/\mathcal{U})e(a/\mathcal{U})\}, \quad \text{if } a \in W_{n+1} \setminus W_n, \end{aligned}$$

after checking that if $a \in W_{n+1}$ and $(b/\mathcal{U})e(a/\mathcal{U})$ then $b \in W_n$.

An easy induction shows that $\hat{}$ maps $W_n$ into $V_n({}^*\mathbb{R})$.

(3) Define the mapping $* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ by ${}^*c = \hat{\mathbf{c}}$ for $c \in V(\mathbb{R})$. It is clear that ${}^*V_n(\mathbb{R}) = \{\hat{a} : a \in W_n\} \subseteq V_n({}^*\mathbb{R})$; so this is consistent with the definition of ${}^*\mathbb{R}$ in (1) above.

To verify Axiom II(b) (Transfer principle) establish the following.

LEMMA. *Let $\phi(x_1, x_2, \dots, x_k)$ be a bounded quantifier statement and let $a_1, \dots, a_k \in W_\infty$. Then the following are equivalent.*

(a) $\phi(\hat{a}_1, \dots, \hat{a}_k)$ *holds in* $V({}^*\mathbb{R})$,

(b) $\phi(a_1(i), \dots, a_k(i))$ *holds in* $V(\mathbb{R})$, a.c.

*Proof.* The equivalence is established by a routine induction on the length of $\phi$. To start the induction off we need the following for $a, b \in W_\infty$.

(i) $\hat{a} \in \hat{b} \Leftrightarrow a(i) \in b(i)$ a.c.,

(ii) $\hat{a} = \hat{b} \Leftrightarrow a(i) = b(i)$ a.c.

The definition of $\hat{}$ gives (i); prove (ii) for $a, b \in W_n$ by induction on $n$.

Setting $a_1, \dots, a_k = \mathbf{c}_1, \dots, \mathbf{c}_k$ (for $c_1, \dots, c_k \in V(\mathbb{R})$) in this lemma gives the transfer principle.

For Axiom III ($\omega_1$-saturation) note first that internal objects are precisely the objects $\hat{a}$ for $a \in W_\infty$. Suppose then that $\hat{a}_1 \supseteq \hat{a}_2 \supseteq \dots$ with each $\hat{a}_k \neq \varnothing$. Taking $I = \mathbb{N}$ to help the notation, choose $b \in W_\infty$ such that for each $i$

$$b(i) \in a_1(i) \cap \dots \cap a_{k_i}(i)$$

where $k_i \leqslant i$ is as large as possible making the set on the right non-empty; $k_i$ and hence $b(i)$ *are* defined a.c. because $a_1(i) \neq \varnothing$ a.c. by hypothesis. Now fix $m$; we will show that $\hat{b} \in \hat{a}_m$. By assumption $a_1(i) \cap \dots \cap a_m(i) \neq \varnothing$ a.c. Moreover, $i \geqslant m$ a.c.; so $k_i \geqslant m$ a.c. and hence $b(i) \in a_m(i)$ a.c. which means that $\hat{b} \in \hat{a}_m$. This completes the proof of the theorem.

For a nonstandard universe satisfying the enlargement axiom, the construction above is modified by taking a larger index set $I$. Let $I_0 = \mathbb{N}$; for each collection $\mathscr{X}$ from $V(\mathbb{R})$ having the finite intersection property, let $I_{\mathscr{X}} = \{F \subseteq \mathscr{X} : \varnothing \neq F \text{ is finite}\}$. Take as the index set the set

$$I = I_0 \times \prod_{\substack{\mathscr{X} \text{ has} \\ \text{f.i.p.}}} I_{\mathscr{X}}.$$

For each $i \in I$ write $i = (i_0, (i_{\mathscr{X}})_{\mathscr{X} \text{ has f.i.p.}})$. Define $U_i \subseteq I$ by

$$U_i = \{j : j_0 \geqslant i_0 \text{ and } j_{\mathscr{X}} \supseteq i_{\mathscr{X}} \text{ for all } \mathscr{X}\}.$$

The collection $(U_i)_{i\in I}$ has the f.i.p; let $\mathscr{U}$ be an ultrafilter on $I$ extending $(U_i)_{i\in I}$.

Now proceed to construct $V({}^*\mathbb{R})$ and ${}^*V(\mathbb{R})$ as in proof of the theorem above. Transfer holds exactly as before, and $\omega_1$-saturation is established with only a minor modification (replace $i$ by $i_0$ in the definition of $k_i$). For enlargement, consider a given $\mathscr{X}$ with f.i.p. Choose $b \in V(\mathbb{R})^I$ such that for each $i$

$$b(i) \in \bigcap_{X \in i_{\mathscr{X}}} X .$$

Then for each fixed $X$, $b(i) \in X$ whenever $X \in i_{\mathscr{X}}$, which happens a.c.; thus $\hat{b} \in {}^*X$.

For a nonstandard universe satisfying $\kappa$-saturation it is necessary to take a yet larger index set $I$ in the above construction; alternatively, the construction of an enlargement can be iterated $\kappa^+$ times. Specifically, set $\mathbb{R}_0 = \mathbb{R}$ and rename ${}^*\mathbb{R}$ and the mapping $*$ in the above enlargement as $\mathbb{R}_1$ and $f_0$ so that $* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ becomes $f_0 : V(\mathbb{R}_0) \to V(\mathbb{R}_1)$. Define $(\mathbb{R}_\alpha, f_\alpha)_{\alpha \leqslant \kappa^+}$ by iteration: obtain $f_\alpha : V(\mathbb{R}_\alpha) \to V(\mathbb{R}_{\alpha+1})$ from $V(\mathbb{R}_\alpha)$ just as $V(\mathbb{R}_1)$, $f_0$ were constructed from $V(\mathbb{R}_0)$, and at limits $\delta$ take direct limits. Finally set ${}^*\mathbb{R} = \mathbb{R}_{\kappa^+}$ and the mapping $* : V(\mathbb{R}) \to V({}^*\mathbb{R})$ is the appropriate direct limit.

## *References*


[An$_1$] R. M. ANDERSON, 'A non-standard representation for Brownian motion and Itô integration', *Israel J. Math.*, 25 (1976), 15–46.

[An$_2$] R. M. ANDERSON, 'Star-finite representations of measure spaces', *Trans. Amer. Math. Soc.*, 271 (1982), 667–687.

[Ar$_1$] L. ARKERYD, 'A nonstandard approach to the Boltzmann equation', *Arch. Rational Mech. Anal.*, 77 (1981), 1–10.

[Ar$_2$] L. ARKERYD, 'A time-wise approximated Boltzmann equation', *IMA Journal of Applied Mathematics*, 27 (1981), 373–383.

[Ar$_3$] L. ARKERYD, 'Intermolecular forces of infinite range and the Boltzmann equation', *Arch. Rational Mech. Anal.*, 77 (1981), 11–21.

[Ar$_4$] L. ARKERYD, 'Asymptotic behaviour of the Boltzmann equation with infinite range forces', *Comm. Math. Phys.*, 86 (1982), 475–484.

[AFHL] S. ALBEVERIO, J. E. FENSTAD, R. HÖEGH-KROHN and T. L. LINDSTRØM, *Nonstandard methods in stochastic analysis and mathematical physics* (no definitive title as yet), Springer-Verlag, to appear.

[AR] R. M. ANDERSON and S. RASHID, 'A nonstandard characterisation of weak convergence', *Proc. Amer. Math. Soc.*, 69 (1978), 327–332.

[Ba] M. T. BARLOW, 'One dimensional stochastic differential equations with no strong solution', *J. London Math. Soc.*, 26 (1982), 335–347.

[Be] A. R. BERNSTEIN, 'A non-standard integration theory for unbounded functions', *Z. Math. Logik Grundlag. Math.*, 20 (1974), 97–108.

[BL] A. R. BERNSTEIN and P. A. LOEB, 'A non-standard integration theory for unbounded functions', *Victoria symposium on nonstandard analysis*, Lecture Notes in Mathematics 369 (Eds. A. E. Hurd and P. A. Loeb, Springer-Verlag, Berlin, 1974), pp. 40–49.

[BS] J. M. BAYOD and K. D. STROYAN, *Foundations of infinitesimal stochastic analysis* (North-Holland) to appear.

[BW] A. R. BERNSTEIN and F. WATTENBERG, 'Nonstandard measure theory', *Applications of model theory to algebra, analysis and probability* (Ed. W. A. J. Luxemburg, Holt, Rinehart and Winston, New York, 1969), pp. 171–185.

[C$_1$] N. J. CUTLAND, 'On the existence of solutions to stochastic differential equations on Loeb spaces', *Z. Wahrsch. verw. Gebiete*, 60 (1982), 335–357.

[C$_2$] N. J. CUTLAND, 'Internal controls and relaxed controls', *J. London Math. Soc.*, 27 (1983), 130–140.

[C$_3$] N. J. CUTLAND, 'Optimal controls for partially observed stochastic systems: an infinitesimal approach', *Stochastics*, 8 (1983), 239–257.

[$C_4$] N. J. CUTLAND, 'Optimal controls for partially observed stochastic systems using nonstandard analysis', *Stochastic differential systems*, Lecture Notes in Control and Information Sciences, 43 (Eds. M. Kohlmann and N. Christopeit, Springer, Berlin, 1982), pp. 276–284.

[D] M. DAVIS, *Applied nonstandard analysis* (Wiley, New York, 1977).

[F] J. E. FENSTAD, 'Is nonstandard analysis relevant for the philosophy of mathematics?', *Synthese*, 1983, to appear.

[GP] P. GREENWOOD and E. PERKINS, 'A conditional limit theorem for random walk, and Brownian local time on square root boundaries', *Ann. Probab.*, 11 (1983), 227–261.

[$He_1$] C. W. HENSON, 'On the nonstandard representation of measures', *Trans. Amer. Math. Soc.*, 172 (1972), 437–446.

[$He_2$] C. W. HENSON, 'Analytic sets, Baire sets and the standard part map', *Canad. J. Math.*, 31 (1979), 663–672.

[$He_3$] C. W. HENSON, 'Unbounded Loeb measures', *Proc. Amer. Math. Soc.*, 74 (1979), 143–150.

[Hl] L. L. HELMS, 'Hyperfinite for spin models', [**$Hu_3$**] (1983), pp. 15–26.

[$Hu_1$] A. E. HURD, 'Nonstandard analysis and lattice statistical mechanics: a variational principle', *Trans. Amer. Math. Soc.*, 263 (1981), 89–110.

[$Hu_2$] A. E. HURD, 'Existence of time evolutions of nonequilibrium states in statistical mechanics', preprint, University of Victoria, B.C., 1981.

[$Hu_3$] A. E. HURD (Editor), *Nonstandard analysis—recent developments*, Lecture Notes in Mathematics 983 (Springer-Verlag, 1983).

[HK] D. N. HOOVER and H. J. KEISLER, 'Adapted probability distributions', preprint, University of Wisconsin, 1982.

[$HL_1$] L. L. HELMS and P. A. LOEB, 'Applications of nonstandard analysis to spin models', *J. Math. Anal. Appl.*, 69 (1979), 341–352.

[$HL_2$] L. L. HELMS and P. A. LOEB, 'Bounds on the oscillation of spin systems', preprint, University of Illinois, Urbana-Champaign.

[$HL_3$] L. L. HELMS and P. A. LOEB, 'A nonstandard proof of the martingale convergence theorem', *Rocky Mountain J. Math.*, 12 (1982), 165–170.

[HP] D. N. HOOVER and E. PERKINS, 'Nonstandard construction of the stochastic integral and applications to stochastic differential equations, I' and 'II', *Trans. Amer. Math. Soc.*, to appear.

[HW] C. W. HENSON and F. WATTENBERG, 'Egoroff's theorem and the distribution of standard points in a nonstandard model', *Proc. Amer. Math. Soc.*, 81 (1981), 455–461.

[HY] E. HEWITT and K. YOSIDA, 'Finitely additive measures', *Trans. Amer. Math. Soc.*, 72 (1952), 46–66.

[Ka] T. KAMAE, 'A simple proof of the ergodic theorem using nonstandard analysis', *Israel J. Math.*, 42 (1982), 284–290.

[$Ke_1$] H. J. KEISLER, 'Hyperfinite model theory', *Logic Colloquium* '76 (Eds. R. O. Gandy and J. M. E. Hyland, North-Holland, Amsterdam, 1977), pp. 5–110.

[$Ke_2$] H. J. KEISLER, *Foundations of infinitesimal calculus* (Prindle, Weber and Schmidt, 1976).

[$Ke_3$] H. J. KEISLER, 'An infinitesimal approach to stochastic analysis', *Mem. Amer. Math. Soc.*, to appear.

[$Ke_4$] H. J. KEISLER, 'Probability quantifiers', to appear in *Applied model theory and strong logics* (Eds. J. Barwise and S. Feferman, Springer-Verlag).

[$Ke_5$] H. J. KEISLER, 'A price adjustment model with infinitesimal traders', preliminary report, University of Wisconsin, 1979.

[$Ke_6$] H. J. KEISLER, 'Nonstandard probability theory and probability logic', unpublished lecture notes, University of Wisconsin, 1979.

[$Ko_1$] S. A. KOSCIUK, 'Stochastic solutions to partial differential equations', [**$Hu_3$**] (1983), pp. 113–119.

[$Ko_2$] S. A. KOSCIUK, 'Nonstandard stochastic methods in diffusion theory', Ph.D. Thesis, University of Wisconsin, Madison, 1982.

[La] G. F. LAWLER, 'A self-avoiding random walk', *Duke Math. J.*, 47 (1980), 655–693.

[$Li_1$] T. L. LINDSTRØM, 'Hyperfinite stochastic integration I', 'II', 'III' and 'addendum', *Math. Scand.*, 46 (1980), 265–292, 293–314, 315–331, 332–333.

[$Li_2$] T. L. LINDSTRØM, 'A Loeb-measure approach to theorems by Prohorov, Sazanov and Gross', *Trans. Amer. Math. Soc.*, 269 (1982), 521–534.

[$Li_3$] T. L. LINDSTRØM, 'The structure of hyperfinite stochastic integrals', *Z. Wahrsch. verw. Gebiete*, to appear.

[$Li_4$] T. L. LINDSTRØM, 'Stochastic integration in hyperfinite dimensional linear spaces', [**$Hu_3$**] (1983), pp. 134–161.

[$Lo_1$] P. A. LOEB, 'A nonstandard representation of measurable spaces and $L_\infty$', *Bull. Amer. Math. Soc.*, 77 (1971), 540–544.

[$Lo_2$] P. A. LOEB, 'A nonstandard representation of measurable spaces, $L_\infty$, and $L_\infty^*$', *Contributions to nonstandard analysis* (Eds. W. A. J. Luxemburg and A. Robinson, North-Holland, Amsterdam, 1972), pp. 65–80.

[$Lo_3$] P. A. LOEB, 'A nonstandard representation of Borel measures and $\sigma$-finite measures', *Victoria symposium on nonstandard analysis*, Lecture Notes in Mathematics 369 (Eds. A. E. Hurd and P. A. Loeb, Springer-Verlag, 1974), pp. 144–152.

[$Lo_4$] P. A. LOEB, 'Conversion from nonstandard to standard measure spaces and applications in probability theory', *Trans. Amer. Math. Soc.*, 211 (1975), 113–122.

[$Lo_5$] P. A. LOEB, 'Applications of nonstandard analysis to ideal boundaries in potential theory', *Israel J. Math.*, 25 (1976), 154–187.

[$Lo_6$] P. A. LOEB, 'A generalisation of the Riesz–Herglotz theorem on representing measures', *Proc. Amer. Math. Soc.*, 71 (1978), 65–68.

[$Lo_7$] P. A. LOEB, 'Weak limits of measures and the standard part map', *Proc. Amer. Math. Soc.*, 77 (1979), 128–135.

[$Lo_8$] P. A. LOEB, 'An introduction to nonstandard analysis and hyperfinite probability theory', *Probabilistic analysis and related topics*, *Vol.* 2 (Ed. A. T. Bharucha-Reid, Academic Press, New York, 1979), pp. 105–142.

[$Lo_9$] P. A. LOEB, 'A regular metrizable boundary for solutions of elliptic and parabolic differential equations', *Math. Ann.*, 251 (1980), 43–50.

[$Lo_{10}$] P. A. LOEB, 'A construction of representing measures for elliptic and parabolic differential equations', *Math. Ann.*, 260 (1982), 51–56.

[$Lo_{11}$] P. A. LOEB, 'A functional approach to nonstandard measure theory', to appear in the proceedings of the S. Kakutani retirement conference.

[$Lo_{12}$] P. A. LOEB, 'Measure spaces in nonstandard models underlying standard stochastic processes', manuscript of address to the International Congress of Mathematicians, Warsaw, 1983.

[Pa] R. L. PANETTA, 'Hyperreal probability spaces: some applications of the Loeb construction', Ph.D. Thesis, University of Wisconsin, 1978.

[$Pe_1$] E. PERKINS, 'A global intrinsic characterisation of Brownian local time', *Ann. Probab.*, 9 (1981), 800–817.

[$Pe_2$] E. PERKINS, 'The exact Hausdorff measure of the level sets of Brownian motion', *Z. Wahrsch. verw. Gebiete*, 58 (1981), 373–388.

[$Pe_3$] E. PERKINS, 'Weak invariance principles for local time', *Z. Wahrsch. verw. Gebiete*, 60 (1982), 437–451.

[$Pe_4$] E. PERKINS, 'On the construction and distribution of a local martingale with a given absolute value', *Trans. Amer. Math. Soc.*, 271 (1982), 261–281.

[$Pe_5$] E. PERKINS, 'Stochastic processes and nonstandard analysis', **[$Hu_3$]** (1983), pp. 162–185.

[PP] R. PARIKH and M. PARNES, 'Conditional probabilities and uniform sets', *Victoria symposium on nonstandard analysis*, Lecture Notes in Mathematics 369 (Eds. A. E. Hurd and P. A. Loeb, Springer-Verlag, 1974), pp. 180–194.

[Rb] A. ROBINSON, *Non-standard analysis* (North-Holland, Amsterdam, 1966).

[$Rd_1$] H. RODENHAUSEN, 'A characterisation of nonstandard liftings of measurable functions and stochastic processes', *Israel J. Math.*, 43 (1982), 1–22.

[$Rd_2$] H. RODENHAUSEN, 'The completeness theorem for adapted probability logic', Ph.D. Thesis, University of Heidelberg, 1982.

[Ry] H. L. ROYDEN, *Real analysis* (Macmillan, New York, 1968).

[RZ] A. ROBINSON and E. ZAKON, 'A set theoretical characterisation of enlargements', *Applications of model theory to algebra, analysis and probability* (Ed. W. A. J. Luxemburg, Holt, Rinehart and Winston, New York, 1969), pp. 109–122.

[S] T. Q. SIBLEY, 'The theory of finitely additive probability using nonstandard analysis', Ph.D. Thesis, Boston University, 1973.

[St] A. STOLL, 'A nonstandard construction of Lévy Brownian motion with applications to invariance principles', Diplomarbeit, Universität Freiburg, 1982.

[SL] K. D. STROYAN and W. A. J. LUXEMBURG, *Introduction to the theory of infinitesimals* (Academic Press, 1976).

[T] D. TALL, 'Elementary axioms and pictures for infinitesimal calculus', *Bull. Inst. Math. Appl.*, 18 (1982), 43–48.

[$W_1$] F. WATTENBERG, 'Nonstandard measure theory—Hausdorff measure', *Proc. Amer. Math. Soc.*, 65 (1977), 326–331.

[$W_2$] F. WATTENBERG, 'Nonstandard measure theory: avoiding pathological sets', *Trans. Amer. Math. Soc.*, 250 (1979), 357–368.

Department of Pure Mathematics,
University of Hull,
22–24 Newland Park,
Hull HU5 2DW,
England.