Applications of non-standard analysis to relativistic quantum mechanics. I

1981 J. Phys. A: Math. Gen. 14 2539

(http://iopscience.iop.org/0305-4470/14/10/010)

J. Phys. A: Math. Gen. 14 (1981) 2539–2551. Printed in Great Britain

Applications of non-standard analysis to relativistic quantum mechanics: I†

C E Francis

Department of Mathematics, Birkbeck College, Malet Street, London WC1E7HX, England

Received 22 July 1980, in final form 16 March 1981

Abstract. Some representations of the Dirac δ function are considered including a new representation. A new theory of Fourier transforms is developed which is better suited to use in physics than the standard theory. The work is of general interest as well as of relevance to subsequent articles. We then give a brief outline of the construction of field theory from quantum mechanics as facilitated by non-standard analysis and a theorem which enables the calculation of a cross section from plane wave states.

1. Introduction

Non-standard analysis is now accepted as a powerful method of proof in mathematics, and its application to quantum physics has often been suggested in view of the severe divergence problems of this subject. Work in this area has already been done by Kelemen and Robinson (1972), Faroukh (1975), Blanchard and Tarski (1978) and Tarski (1978). The present approach as outlined in § 3 is entirely unrelated to these.

The foundational work on non-standard analysis is Robinson (1966). Introductory accounts can be found in Luxembourg (1973), Voros (1974), and Kelemen and Robinson (1972). An excellent full account is given by Davis (1977). Machover and Hirschfield (1969) give an account with a clear insight into the construction of non-standard analysis. The purpose of the present paper is to develop some mathematical methods using non-standard analysis with a view to applying them to develop a mathematical form of quantum field theory. We describe this briefly in § 3.1. The detail is given in subsequent papers. In § 3.2 we give an example to illustrate the power of the new method.

It may be thought at first sight that the results of § 2 are not really new and the mathematical difficulties have just been transferred to non-uniqueness problems. In fact for all practical purposes non-uniqueness is not a problem. (Quantum mechanics already presents an example of this fact.) Further there is a fundamental new feature in that the new theory enables the use of δ functions and Fourier analysis on a wider class of functions. Physicists work on this wider class anyway as an algorithm which they know works from experience. However, such an algorithm cannot lead to any explanation or deep understanding of physics unless it can be shown to be mathematically valid.

† This paper will form part of a PhD thesis to be submitted to the University of London. The work was done under the supervision of Professor C S Sharma.

0305-4470/81/102539+13$01.50

For the purpose of understanding the main part of this work an intuitive idea of non-standard analysis should be adequate. It is a theory of infinities and infinitesimals which behave like the reals as understood by Leibniz and which is suitable for the development of the calculus. The outline is as follows.

Starting with a universe U, consisting of a set S of individuals (e.g. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Z}$, a Hilbert space H) together with all relations, relations between relations etc, defined on S we construct a larger universe W, which has the following properties.

To every element r of U corresponds an element $^*r$ of W, Then $^*r$ is called a standard element of W. It is usual in this case to suppress the * when no ambiguity will arise. For example if r is a function $r: S \to S$, say then $^*r: {}^*S \to {}^*S$ and whenever $r(s) = t$ we have $^*r(^*s) = {}^*t$. Then, since we embed S in $^*S$ so that $^*s = s$ and $^*t = t$, $^*r$ is an extension of r.

There is a universe $^*U \subset W$, called the non-standard universe. (The definition of a universe is technical, and $^*U \neq W$ for an infinite set S). The elements of W will be called internal if they are in $^*U$, and external if they are not. In particular every standard element is internal.

We now come to the all important transfer principle: 'any sentence which is true about U may be interpreted as a true sentence about the internal sets $^*U$'. We used this above when we wrote $r(s) = t \Rightarrow {}^*r(^*s) = {}^*t$. This sentence is also true when r is a map between general elements of U.

For example the phrase: '$\mathbb{R}$ is an ordered field' can be re-interpreted '$^*\mathbb{R}$ is an ordered field', so that we can add and multiply infinite and infinitesimal numbers. But the phrase: '$\mathbb{R}$ contains only finite elements' must be re-interpreted '$^*\mathbb{R}$ contains only $*$ finite elements'. In fact infinite and infinitesimal elements are $*$ finite, and both $\mathbb{R}$ and the set of finite numbers are external. It was the lack of this distinction which caused the trouble in Liebniz's theory. The elements of $^*\mathbb{R}$ will be called hyperreal.

1.1. *Analysis*

We work on $^*\mathbb{R}$. For the purposes of this article we use standard definitions of integrals and derivatives and extend them by the transfer principle to internal functions. The integral is the proper Riemann integral. Let $x, y \in {}^*\mathbb{R}$. We say that $x \approx y$ whenever $x - y$ is infinitesimal. If $y \in \mathbb{R}$ and $x \approx y$ we write $^0x = y$ and say that y is the standard part of x.

We require the non-standard formulation for the limit of a sequence. Note that a sequence s_n is a function $s: \mathbb{Z}^+ \to \mathbb{R}$ so $^*s: {}^*\mathbb{Z}^+ \to {}^*\mathbb{R}$ and we may write $^*s = s$. Then $s_n \to l \in{}'$ iff $s_n \approx l$ for all infinite $n \in {}^*\mathfrak{I}^+$. For example the improper Riemann integral exists provided $^0(\int_{-\mu}^{\nu} \mathrm{d}x f(x))$ exists and is independent of the infinite positive numbers μ, n.

A standard function is continuous at $x \in \mathbb{R}$ iff $f(y) \approx f(x)$ whenever $y \approx x$. An internal function for which $x \approx y \Rightarrow f(x) \approx f(y)$ is called microcontinuous.

1.2. *Some non-standard theorems*

The first two are somewhat modified versions of the infinitesimal prolongation theorem.

Proposition 1.2.1. Let f be an internal function $f: {}^*\mathbb{R} \to {}^*\mathbb{R}$ such that $\forall y_1 \in \mathbb{R}^+$, $y \in {}^*\mathbb{R}$ and $y > y_1 \Rightarrow f(y) \approx 0$. Then $\exists y_0 \in {}^*\mathbb{R}^+$, $y_0 \approx 0$ such that $y > y_0 \Rightarrow f(y) \approx 0$.

Proposition 1.2.2. Let f be an internal function $f: ({}^*\mathbb{R}^+)^n \to {}^*\mathbb{R}$ where $n \in \mathbb{Z}^+$. Let $f = f(m, \ldots, m_n) = f(m)$ be such that $f(m) \approx 0$ whenever each m_i is finite for $i = 1, \ldots, n$. Then $\exists M \in {}^*\mathbb{R}$, infinite such that if $m_i < M$ for each $i = 1, \ldots, n$ then $f(m) \approx 0$.

The third is an (apparently) stronger version of the non-standard criterion for the limit.

Proposition 1.2.3. Let f be a standard function $f: (\mathbb{R}^+)^n \to \mathbb{R}$ where $n \in \mathbb{Z}^+$. Let $f = f(m_1, \ldots, m_2) = f(m)$ be such that $\exists M$ such that if for each $i = 1, \ldots, n$ m_i is infinite and $m_i < M$ then $f(m) \approx a \in \mathbb{R}$. Then $f(m) \approx a$, $\forall$ infinite $m_i \in {}^*\mathbb{R}^+$ $i = 1, \ldots, n$ i.e.

$$\lim_{\substack{m_i \to \infty \\ \text{each } i}} f(m) = a.$$

The proofs follow by obvious modifications of the ones given by e.g. Davis (1977) and are given in the Appendix.

2. Some basic mathematical results

2.1. *The Riemann–Lebesgue lemma*

For the standard version and proof the reader is referred to any standard work on mathematical methods of physics e.g. Olver (1974). The non-standard form which we shall use is: If $q(t)$ is standard, continuous on the interval $[a, \infty) \subseteq \mathbb{R}$ and $\int_\mu^\nu \mathrm{d}t q(t)\, \mathrm{e}^{\mathrm{i}xt} \approx 0$ for all hyperreal x such that $|x| > x_0 \in \mathbb{R}$ and for all infinite positive μ, ν (this is equivalent to uniform convergence of the integral) then for finite a, $\int_a^\mu \mathrm{d}t q(t)\, \mathrm{e}^{\mathrm{i}xt} \approx 0$ for infinite positive μ, $|x|$.

Note. Many books require $\int_0^\infty \mathrm{d}t q(t)\, \mathrm{e}^{\mathrm{i}xt}$ to be absolutely convergent, but in fact uniform convergence is adequate, so that the Riemann–Lebesgue lemma holds for a wide variety of functions such as $q(t) = 1$, $q(t) = 1/t$. However, it is difficult to write down conditions on q to ensure uniform convergence without creating an undesirable restriction.

2.2. *The non-standard treatment of the Dirac δ function*

Non-standard treatments of distribution can be found in Robinson (1966) as well as in Luxemburg (1962) and Stroyan and Luxemburg (1976). However, we believe that once armed with the methods of non-standard analysis a physicist has no need of distributions. For any distribution D we can write down any number of representations of D which are internal functions. Then we need study only the representation, which is considerably simpler than the study of the distribution. We shall suppress the word representation. For example for the Dirac δ function we may write as

$$\delta(x) \approx_T \begin{cases} 0 & \text{if } |x| < (2\nu)^{-1} \\ \nu & \text{if } |x| \geqslant (2\nu)^{-1} \end{cases} \qquad \text{where } 0 < \nu \text{ is infinite} \tag{1}$$

$$\delta(x) \approx_T \sqrt{\nu/\pi}\, \mathrm{e}^{-\nu x^2} \qquad \text{where } 0 < \nu \text{ is infinite} \tag{2}$$

$$\delta(x) \approx_T \begin{cases} (\pi\nu\chi)^{-1} \sin \nu^2 x & \text{for } x \neq 0 \\ \nu/\pi & \text{for } x = 0 \end{cases} \qquad \text{where } 0 < \nu \text{ is infinite.} \tag{3}$$

We use the symbol $\approx_T$ in order to indicate the need for a suitable test space T of functions f such that the δ function has the required behaviour:

$$\int_a^b \mathrm{d}x\, \delta(x-y)f(x) \approx f(y) \qquad \text{for } a<y<b.$$

The test space depends on the choice of δ function. For (1) we need only f to be micro continuous, but for (2) and (3) we also require boundedness conditions as $x \to \infty$. We require further conditions if we wish to use derivatives of the δ function:

$$\delta'(x) \approx_T -(2\nu^{3/2}x/\sqrt{\pi})\, \mathrm{e}^{-\nu x^2} \tag{2a}$$

$$\delta'(x) \approx_T \begin{cases} (\nu/\pi x)\cos\gamma^2 x - (\pi x^2\nu)^{-1}\sin\gamma^2 x & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases} \tag{3a}$$

(2a) is infinitesimal for finite x, (3a) is frequently infinite. This is a new feature of the δ function and we need to understand what it means. Consider

$$\int_{\mu_1}^{\mu_2} \mathrm{d}x\, \delta'(x)f(x) = [\delta(x)f(x)]_{\mu_1}^{\mu_2} - \int_{\mu_1}^{\mu_2} \mathrm{d}x\, \delta(x)f'(x) \qquad \mu_1 < 0 < \mu_2.$$

To obtain the required behaviour we need

(i) $\delta(\mu_j)f(\mu_j) \approx 0$, $j = 1, 2$,

(ii) f' is in the test space T for the relevant δ function.

We shall apply the conditions defining T in a fairly *ad hoc* manner in order to suit the particular problems that we shall have, but we observe that there is no need to restrict ourselves to the test space of infinite differentiable functions of rapid decay.

2.3. Fourier transforms

Just as it is convenient nct to have a unique δ function it is also convenient not to have a unique Fourier transform. We define, for positive hyperreals μ, ν and for the internal function $f(x)$, $*$-integrable on any interval

$$F_\mu^\nu f(t) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\mu}^{\nu} \mathrm{d}x\, f(x)\, \mathrm{e}^{-\mathrm{i}xt}$$

unless otherwise stated μ, ν will be taken to be infinite.

If for any infinite μ, ν, $F_\mu^\nu f(t)$ is microcontinuous, near standard and ${}^0(F_\mu^\nu f(t))$ is independent of μ, ν at $t \in \mathbb{R}$ we can define $Ff(t) \equiv {}^0(F_\mu^\nu f(t))$, the standard Fourier transform. If Ff can be defined on $\mathbb{R}$ and $F_\mu^\nu f$ is microcontinuous on ${}^*\mathbb{R}$ then the approximation theorem (Davis 1977) states that Ff is continuous on $\mathbb{R}$ and that $Ff(t) \approx F_\mu^\nu f(t)$ $\forall t \in {}^*\mathbb{R}$.

Proposition 2.3.1. If f is standard and $\int_{-\infty}^{\infty} \mathrm{d}x\, |f(x)|$ exists we can define Ff, continuous on $\mathbb{R}$ and $Ff(t) \approx F_\mu^\nu f(t)$ $\forall t \in {}^*\mathbb{R}$.

Proof. We have only to prove that $F_\mu^\nu f$ is microcontinuous on ${}^*\mathbb{R}$, i.e. that $\forall t \in {}^*\mathbb{R}$, $\mathrm{d}t \approx 0$ $|F_\mu^\nu f(t+\mathrm{d}t) - F_\mu^\nu f(t)| \approx 0$. Now if $\mathrm{d}t \approx 0$, then by the infinitesimal prolongation theorem 1.2.2 $\exists M$, infinite such that $\mu, \nu < M \Rightarrow \mu\, \mathrm{d}t \approx \nu\, \mathrm{d}t \approx 0$. Then $\forall \mu, \nu < M$

$$|F_\mu^\nu f(t+\mathrm{d}t) - F_\mu^\nu f(t)| \leq \int_{-\mu}^{\nu} \mathrm{d}x\, |f(x)|\, |\mathrm{e}^{-\mathrm{i}x\,\mathrm{d}t} - 1|.$$

But

$$e^{-ix\,dt} \approx 1 \qquad \text{for } -\mu \leqslant x \leqslant \nu$$

so $F^{\nu}_{\mu} f$ is microcontinuous on ${}^*\mathbb{R}$ for $\mu, \nu < M$. Thus Ff is continuous on $\mathbb{R}$ and

$$Ff(t) \approx (F^{\mu}_{\nu} f)(t) \qquad \text{for } t \in {}^*\mathbb{R} \qquad \mu, \nu < M.$$

Therefore by 1.2.3

$$\lim_{\substack{t\to\infty \\ \mu,\nu\to\infty}} (F^{\mu}_{\nu} f(t) - Ff(t)) = 0$$

$$\Rightarrow Ff(t) \approx F^{\nu}_{\mu} f(t) \qquad \text{for } t \in {}^*\mathbb{R}, \forall \text{ infinite } \mu, \nu.$$

The non-standard Fourier transform has a much wider application than the standard one. For example if $f(x) = 1$ then

$$(F^{\nu}_{\mu} 1)(t) = \left(\frac{1}{2\pi}\right)^{1/2} \frac{\mathrm{i}}{t}(e^{-it\nu} - e^{it\mu})$$

which is a δ function as we shall show.

We first define a test space as follows. Let T be the set of differentiable standard functions $\mathbb{R} \to \mathbb{R}$ such that for $f \in T$ $\int_{\mu_1}^{\mu_2} dt\, e^{ixt} f(t)/t \approx 0$ for all hyperreal x with $|x| > x_0 \in \mathbb{R}$ and for all infinite positive and all infinite negative μ_1, μ_2.

Lemma 2.3.2. Let a be finite, $f \in T$. Then $\int_{\mu_1}^{\mu_2} dt\, e^{ixt} f(t)/(t-a) \approx 0$ with μ_1, μ_2 as in the preceding definition.

Proof.

$$\int_{\mu_1}^{\mu_2} dt\, e^{ixt} \frac{f(t)}{t-a} = \int_{\mu_1}^{\mu_2} dt\, e^{ixt} \frac{f(t)}{t}\left(1 + \frac{a}{t-a}\right)$$

$$\approx a \int_{\mu_1}^{\mu_2} dt\, e^{ixt} \frac{f(t)}{t} \frac{1}{t-a}$$

$$= \frac{a}{\eta - a} \int_{\mu_1}^{\mu_2} dt\, e^{ixt} \frac{f(t)}{t}$$

(for some $\mu_1 \leqslant \eta \leqslant \mu_2$ by the mean value theorem)

$$\approx 0.$$

Theorem 2.3.3. Let $f \in T$. Then

$$I \equiv \int_{-\kappa}^{\lambda} dt \frac{\mathrm{i}}{t}(e^{-it\nu} - e^{it\mu}) f(t+a) \approx 2\pi f(a)$$

where μ, ν are positive infinite, a is finite and $-\kappa < a < \lambda$.

Remark. We may without loss of generality set $a = 0$.

Lemma 2.3.4. Let f be differentiable at 0. Then for infinitesimal $\eta > 0$

$$\int_{-\eta}^{\eta} dt\, f(t) \frac{\mathrm{i}}{t}(e^{-it\nu} - e^{it\mu}) \approx \int_{-\eta}^{\eta} dt\, f(0) \frac{\mathrm{i}}{t}(e^{-it\nu} - e^{it\mu})$$

Proof of lemma. Since f is differentiable at 0 we have from Davis (1977 pp 63–6), for infinitesimal t, $f(t)=f(0)+(f'(0)+\alpha)t$ where α is infinitesimal. Then

$$\int_{-\eta}^{\eta} \mathrm{d}t\, f(t)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}+\mathrm{e}^{\mathrm{i}t\mu})$$

$$=\int_{-\eta}^{\eta} \mathrm{d}t \left\{ f(0)\frac{\mathrm{i}}{t}(e^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})+(f'(0)+\alpha)(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})\right\}$$

$$\approx\int_{-\eta}^{\eta} \mathrm{d}t\, f(0)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})$$

as required.

Proof of theorem. By the RL lemma $\forall y\in\mathbb{R}^+$ and $\kappa,\lambda\in{}^*\mathbb{R}^+$, $y<\kappa,\lambda$

$$\left(\int_{-\kappa}^{-y}+\int_{y}^{\lambda}\right)\mathrm{d}t\, f(t)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})\approx 0.$$

Hence by the infinitesimal prolongation theorem 1.2.1 $\exists\eta_0\approx 0$ such that

$$\eta>\eta_0\Rightarrow\left(\int_{-\kappa}^{-\eta}+\int_{\eta}^{\lambda}\right)\mathrm{d}t\, f(t)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})\approx 0$$

thus for infinitesimal $\eta>\eta_0$

$$I\equiv\int_{-\kappa}^{\lambda}\mathrm{d}t\, f(t)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})$$

$$\approx\int_{-\eta}^{\eta}\mathrm{d}t\, f(0)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})$$

by 2.3.4. Now for $\xi\in\mathbb{R}^+$ and $\tau>0$, infinite we have

$$\int_{\eta}^{\tau}\mathrm{d}t\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})\approx 0$$

by the RL lemma. So by the infinitesimal prolongation theorem 1.2.1 $\exists\xi_0\approx 0$ such that

$$\xi>\xi_0\Rightarrow\int_{\xi}^{\tau}\mathrm{d}t\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})\approx 0.$$

Then taking $0\approx\eta>\max(\xi_0,\eta_0)$ we obtain

$$I\approx\int_{-\eta}^{\eta}\mathrm{d}t\, f(0)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})$$

$$\approx\int_{-\tau}^{\tau}\mathrm{d}t\, f(0)\frac{\mathrm{i}}{t}(\mathrm{e}^{-\mathrm{i}t\nu}-\mathrm{e}^{\mathrm{i}t\mu})$$

$$=\int_{-\tau}^{\tau}\mathrm{d}t\, f(0)\frac{\mathrm{i}}{t}[(\mathrm{e}^{-\mathrm{i}t\nu}-1)-(\mathrm{e}^{\mathrm{i}t\mu}-1)].$$

There is a removable singularity at $\tau=0$. We draw the contour of figure 1. Then

$$\int_{-\tau}^{\tau}\mathrm{d}t\frac{\mathrm{i}}{t}(\mathrm{e}^{\mathrm{i}t\mu}-1)=-\int_{\mathrm{C}_1}\mathrm{d}t\frac{\mathrm{i}}{t}(\mathrm{e}^{\mathrm{i}t\mu}-1).$$

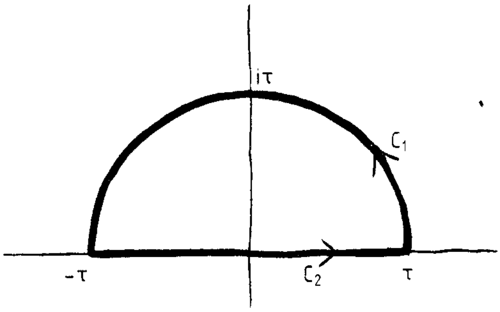

Figure 1. Contour for the evaluation of $\int_{-\tau}^{\tau} \mathrm{d}t\ \mathrm{i}(\mathrm{e}^{\mathrm{i}t\mu}-1)/t$.

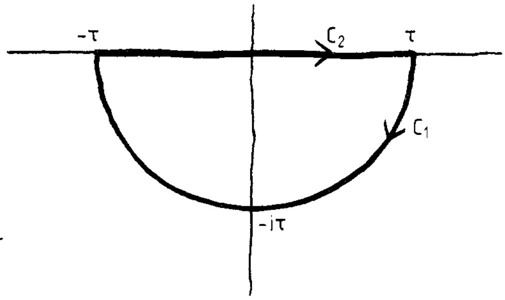

Figure 2. Contour for the evaluation of $\int_{-\tau}^{\tau} \mathrm{d}t\ \mathrm{i}(\mathrm{e}^{-\mathrm{i}t\nu}-1)/t$.

Now (as in Jordan's lemma)

$$\left| \int_{C_1} \mathrm{d}t \frac{\mathrm{e}^{\mathrm{i}t\mu}}{t} \right| \leqslant \int_0^{\pi} \mathrm{d}\theta \exp(-\tau\mu \sin\theta)$$

$$= 2\int_0^{\pi/2} \mathrm{d}\theta \exp(-\tau\mu \sin\theta)$$

$$\leqslant 2\int_0^{\pi/2} \mathrm{d}\theta \exp(-\tau\mu\theta/\pi)$$

$$= \frac{2\pi}{\mu\nu}[\exp(-\tau\mu\theta/\pi)]_0^{\pi/2}$$

$$\approx 0.$$

So

$$\int_{C_1} \mathrm{d}t \frac{\mathrm{i}}{t}(\mathrm{e}^{\mathrm{i}t\mu}-1) \approx \pi.$$

Similarly using the contour of figure 2 we obtain

$$\int_{-\tau}^{\tau} \mathrm{d}t\ (\mathrm{e}^{-\mathrm{i}t\nu}-1)\frac{\mathrm{i}}{t} \approx \pi$$

thus $I \approx 2\pi f(0)$ as required.

We therefore have

$$(F_\mu^\nu 1)(t) \approx_T \sqrt{2\pi}\delta(t).$$

We can now develop a theory of Fourier transforms. This is much more general than the standard theory. We proceed by examining some familiar results in the new theory. It is not difficult to generalise T to include functions with a finite number of discontinuities, provided we can define a left and a right derivative at each point in $\mathbb{R}$.

2.4. *Some miscellaneous results*

Theorem 2.4.1. (Inversion formula). Let $f \in T$, and Z be a finite hyperreal. Then

(i) for infinite $\kappa, \lambda, \mu, \nu$ $F_\kappa^\lambda F_\mu^\nu f(z) \approx f(-z)$,

(ii) if in addition Ff exists then FFf exists and $FFf(z) = f(-z)$.

Proof. (i)

$$F_\kappa^\lambda F_\mu^\nu f(z) = \frac{1}{2\pi}\int_{-\kappa}^{\lambda} dt\, e^{-izt}\int_{-\mu}^{\nu} dx\, e^{ixt} f(x)$$

$$= \frac{1}{2\pi}\int_{-\mu}^{\nu} dx\, f(x)(F_\kappa^\lambda 1)(z+x)$$

(since we may interchange the order of integration for the proper Riemann integral)

$f(-z) \approx$ since $f \in T$.

(ii) Now for any finite κ, λ and infinite μ, ν

$$F_\kappa^\lambda Ff(z) \approx F_\kappa^\lambda F_\mu^\nu f(z)$$

So by 1.2.2 $\exists M$ infinite such that for any infinite $\kappa, \lambda < M$

$$F_\kappa^\lambda Ff(z) \approx F_\kappa^\lambda F_\mu^\nu f(z) \qquad \text{where } \mu, \nu \text{ are infinite}$$

$$\approx f(-z) \qquad \text{by (i).}$$

But then 1.2.3 establishes that this holds for all infinite κ, λ, and the continuity of f implies that FFf exists and so $FFf(z) = f(-z)$ as required.

Theorem 2.4.2. (Parseval's theorem). Let $f, g \in T$. Then
(i) for all positive infinite $\kappa, \lambda, \mu, \nu$

$$\int_{-\lambda}^{\kappa} dx\, (F_\mu^\nu g(x))^* F_\mu^\nu f(x) \approx \int_{-\mu}^{\nu} dy\, g^*(y) f(y)$$

(ii) if Ff and Fg exist then

$$\int_{+\infty}^{\infty} dx\, (Fg(x))^* Ff(x) = \int_{-\infty}^{\infty} dy g^*(y) f(y).$$

Proof. (i) Let λ, κ be infinite and let μ, ν be finite. Then

$$\int_{-\lambda}^{\kappa} dx\, (F_\mu^\nu g)^*(x) F_\mu^\nu f(x)$$

$$= \frac{1}{2\pi}\int_{-\lambda}^{\kappa} dx \int_{-\mu}^{\nu} dy \int_{-\mu}^{\nu} dz\, e^{ixy}\, e^{-ixz} g^*(y) f(x)$$

$$= \frac{1}{2\pi}\int_{-\mu}^{\nu} dy \int_{-\mu}^{\nu} dz\, g^*(y) f(z)(F_\lambda^\kappa 1)(z-y)$$

$$\approx \int_{-\mu}^{\nu} dy\, g^*(y) f(y)$$

since μ, ν are finite. By the infinitesimal prolongation theorem 1.2.2 this holds $\forall \mu$, $\nu < M$, where M is an infinite constant. Thus we have for all infinite $\lambda, \kappa, \mu, \nu < M$

$$h(\lambda, \kappa, \mu, \nu) \equiv \int_{-\lambda}^{\kappa} dx\, (F_\mu^\nu g)^*(x) F_\mu^\nu f(x) - \int_{-\mu}^{\nu} dy\, g^*(y) f(y) \approx 0$$

h is a standard function. So 1.2.3 gives (i).

(ii) Now suppose Fg and Ff exist. Then for any finite κ, λ and infinite μ, ν

$$\int_{-\lambda}^{\kappa} \mathrm{d}x\,(Fg)^*(x)Ff(x) \approx \int_{-\lambda}^{\kappa} \mathrm{d}x\,(F_\mu^\nu g)^*(x)F_\mu^\nu f(x).$$

So by 1.2.2 there exists an infinite M such that this is true $\forall\kappa, \lambda < M$. But then 1.2.3 and (i) give, for infinite κ, λ, μ, $\nu > 0$

$$\int_{-\lambda}^{\kappa} \mathrm{d}x\,(Fg)^*(x)Ff(x) \approx \int_{\mu}^{\nu} \mathrm{d}y\, g^*(y)f(y).$$

Since the LHS is independent of μ, ν and the RHS is independant of κ, λ, neither side depends on the infinite constants and we have (ii).

Remark. Parseval's Theorem is obtained by putting $g = f$.

Theorem 2.3.3. (Shifting theorems).

(i) $F_\mu^\nu f(t)\,\mathrm{e}^{\mathrm{i}at})(x) = F_\mu^\nu f(x-a)$.

(ii) $(F_\mu^\nu f(t-a))(x) = \mathrm{e}^{-\mathrm{i}ax}(F_{\mu-a}^{\nu-a} f)(x)$.

Proof. Elementary.

Convolution. Let y, f be internal functions, $*$-integrable on any interval. Let κ, λ be positive hyperreals. Then we define the convolution

$$g_\kappa^\lambda * f(s) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\kappa}^{\lambda} \mathrm{d}t\, g(s-t)f(t).$$

If for any infinite κ, $\lambda g_\kappa^\lambda * f$ is microcontinuous, near standard and ${}^0(g_\kappa^\lambda * f(s))$ is independant of κ, λ at $S \in \mathbb{R}$ we can define $g*f(s) \equiv {}^0(g_\kappa^\lambda * f(s))$. The approximation theorem (Davis 1977) states that if $g*f$ can be defined on $\mathbb{R}$ and $g_\kappa^\lambda * f$ is microcontinuous on ${}^*\mathbb{R}$ then $g*f$ is continuous on $\mathbb{R}$ and $\forall s \in {}^*\mathbb{R}$

$$g*f(s) \approx g_\kappa^\lambda * f(s).$$

Proposition 2.4.4. Let f, g be standard and such that g is continuous and $\int_{-\infty}^{\infty} |f|\,\mathrm{d}x$ exists. Then we can define $g*f$. $g*f$ is continuous on $\mathbb{R}$ and $\forall s \in {}^*\mathbb{R}$, and infinite κ, λ $g*f(s) \approx g_\kappa^\lambda * f(s)$. The proof is very similar to that of proposition 2.3.1 and is omitted.

Theorem 2.4.5. (The convolution theorem.) If Ff and Fg exist and κ, λ are infinite then

$$Ff(x)Fg(x) = F(g_\kappa^\lambda * f)(x).$$

Proof. Let κ, λ be finite. Then for finite x, infinite μ, ν

$$F_\kappa^\lambda f(x)Fg(x) \approx \left(\frac{1}{2\pi}\right)^{1/2} \int_{-\kappa}^{\lambda} \mathrm{d}t\, F_\mu^\nu\, g(x)\,\mathrm{e}^{-\mathrm{i}xt} f(t)$$

$$= \left(\frac{1}{2\pi}\right)^{1/2} \int_{-\kappa}^{\lambda} \mathrm{d}t (F_\mu^\nu g(s-t)(x)f(t)$$

by the shifting theorem and since Fg exists

$$=F^{\nu}_{\mu}(g^{\lambda}_{\kappa}*f)(x).$$

Thus by 1.2.2 there exists an infinite M such that $\forall$ infinite κ, $\lambda<M$

$$\begin{aligned} Ff(x)Fg(x) &\approx F^{\lambda}_{\kappa}f(x)Fg(x) \\ &\approx F^{\nu}_{\mu}(g^{\lambda}_{\kappa}*f)(x) \qquad \text{for any infinite } \mu, \nu>0 \\ &\approx F(g^{\lambda}_{\kappa}*f)(x) \end{aligned}$$

since the LHS is continuous and near standard for finite x and independent of μ, ν. Both sides are standard so we have equality. Curiously it is not necessary to be able to define $g*f$.

3. Quantum mechanics

3.1. *Quantum field theory*

The object of the present paper is to develop mathematics for use in the construction of field theory from quantum mechanics. Here we outline the procedure as a means of motivating the present paper.

The wavefunctions of one-particle states are square integrable functions. We replace the improper integral in the standard inner product by a proper integral on a fixed infinite segment of the hyperreal line including the reals. This inner product is identical up to an infinitesimal to the standard one. In other words we make an external modification to the standard quantum mechanical model. The new model will make the same predictions as the old one since the coordinate space wavefunctions and rules for time development are unchanged and we take the standard part of the inner product when calculating a probability.

We embed the new model in a larger one including plane waves in the state space. Then, using non-standard analysis as described here we construct field theory on the larger model. It is to be noted that this construction would not be possible without making the external modification referred to in the preceding paragraph.

We define non-interacting multiparticle state spaces as the direct product of single particle state spaces, and Fock space as the direct sum of spaces with different numbers of particles. Since we are working on a space including plane waves we can immediately define creation and annihilation operators. In the standard theory this can only be done hypothetically. We are then able to define field operators in terms of a proper non-standard integral of creation and annihilation operators.

The theory develops along intuitively compelling lines. Relativity demands that no interaction may be propagated faster than the speed of light and that the operators describing interactions shall be covariant. This leads us, almost inevitably, to the construction of local field operators as building blocks for interaction operators. Unlike the standard theory, the field operators are not modified when interactions are introduced—indeed they could not be.

To carry out the above construction it is necessary to take into account spin and any other internal discrete degrees of freedom, and include antiparticles in the description. To this end we postulate wavefunctions as generalisations of the Dirac and photon wavefunction. In the simplest cases the field operators are essentially similar to the Dirac field and the Gupta–Bleuler form of the electromagnetic field, except that these now become well defined, non-standard entities. We are lead automatically to a compelling form of QED which has electromagnetism as the classical correspondence and the interacting Dirac equation as the semiclassical correspondence. It also appears that we can generalise the theory to a model for hadrons which, if true, would solve the problem of quark confinement, give a single theory for all types of physical particle and should lead to a single theory of interactions (excepting perhaps gravity).

The new theory leads to a modification of the Feynman graph rules which accounts for the divergences occurring therein and justifies regularisation according to a program similar to the one currently used. The divergences of certain Feynman diagrams result from a breakdown in the standard algorithm of the Dirac function. The new theory provides a reason for the divergences and justification for the procedure of subtracting off the infinite part.

A systematic development of quantum field theory along these lines has already been achieved and will be reported in a subsequent paper.

3.2. Calculation of a cross section

In standard quantum mechanics we can calculate a cross section from S-matrix elements using wave packets. The calculation is extremely long and tortuous. It is also possible to 'fudge' a calculation from plane-wave states by using the (meaningless) relation

$$\delta^2(x) = \delta(x)\delta(0)$$

(see e.g. Bjorken and Drell 1964, ch 7). In the non-standard theory essentially the same calculation goes through without fudging by using the following theorem.

Theorem 3.2.1. Let κ, λ be positive infinite reals. Let

$$d(x) = \left(\frac{1}{2\pi}\right)\int_{-\kappa}^{\lambda} \mathrm{d}t\, \mathrm{e}^{-\mathrm{i}xt}$$

$$= \left(\frac{1}{2\pi}\right)\frac{\mathrm{i}}{x}(\mathrm{e}^{-\mathrm{i}\lambda x} - \mathrm{e}^{\mathrm{i}\kappa x}) =_T \delta(x).$$

Then $d^2(x)/d(0) \simeq_T \delta(x)$ where the notation, $\simeq_T$ and the test space, T are as defined § 2 above.

Proof. Let $f \in T$, then it is required to prove that

$$\mathrm{I} \equiv \int_{-\mu}^{\nu} \mathrm{d}x\, f(x)\frac{d^2(x)}{d(0)} \approx f(0)$$

for μ, $\nu > 0$. Now

$$I = -\frac{1}{2\pi}\int_{-\mu}^{\nu} \mathrm{d}x\, f(x)\frac{1}{(\kappa+\lambda)x^2}(\mathrm{e}^{-\mathrm{i}\lambda x} - \mathrm{e}^{\mathrm{i}\kappa x})^2$$

$$\approx -\frac{1}{2\pi}\int_{-\mu}^{\nu} \mathrm{d}x\,\frac{1}{x}\frac{1}{(\kappa+\lambda)}\{f'(x)(\mathrm{e}^{-\mathrm{i}\lambda x} - \mathrm{e}^{\mathrm{i}\kappa x})^2$$

$$+\mathrm{i}f(x)(\kappa-\lambda)\,\mathrm{e}^{\mathrm{i}(\kappa-\lambda)x}(\mathrm{e}^{\mathrm{i}(\kappa+\lambda)x} + \mathrm{e}^{-\mathrm{i}(\kappa+\lambda)x} - 2)$$

$$+\mathrm{i}f(x)(\kappa+\lambda)\,\mathrm{e}^{\mathrm{i}(\kappa-\lambda)x}(\mathrm{e}^{\mathrm{i}(\kappa+\lambda)x} - \mathrm{e}^{\mathrm{i}(\kappa+\lambda)x})\}.$$

The first term is infinitesimal.

If $\kappa-\lambda$ is finite the second term is infinitesimal because of the factor $1/(\kappa+\lambda)$. (True for all sufficiently small infinite μ, ν, and so for all μ, ν.) If $\kappa-\lambda$ is infinite the second term is infinitesimal by the Riemann–Lebesgue lemma. So

$$I \approx \frac{1}{2\pi}\int_{-\mu}^{\nu} \mathrm{d}x\, f(x)\frac{i}{x}(\mathrm{e}^{-2\mathrm{i}\lambda x}-\mathrm{e}^{2\mathrm{i}\kappa x})$$
$$\approx f(0)$$

as required.

Appendix

Proof of 1.2.1. Suppose for all

$$y_1 \in \mathbb{R}^+, y \in {}^*\mathbb{R} \qquad y > y_1 \Rightarrow f(y) \approx 0.$$

Then for all $y_2 \in {}^*\mathbb{R}$ such that $y_2 \approx y_1 \in \mathbb{R}^+$ we have

$$y > y_2 \Rightarrow f(y) < y_2 \dots \tag{A1}$$

because clearly

$$y_2 > \tfrac{1}{2}y_1 \qquad \text{and} \qquad y > \tfrac{1}{2}y_1 \in \mathbb{R}^+ \Rightarrow f(y) \approx 0 < y_2.$$

Let

$$S = \{y_2 \in {}^*(0, 1) | y > y_2 \Rightarrow f(y) < y_2\},$$

S is clearly internal. But (A1)$\Rightarrow S$ contains the set

$${}^*(0, 1) \setminus I$$

where I is the set of infinitesimals, which is external. So S must contain an infinitesimal y_0, which proves the theorem.

Proof of 1.2.2. Let

$$S = \{M \in {}^*\mathbb{R}^+ | m_i \leqslant M, \text{for each } i = 1, \dots, n \Rightarrow m_i f(m) < 1\}$$

S is internal. But S contains the set of finite numbers, F, which is external. So S must contain an infinite M.

Proof of 1.2.3. Given $\varepsilon > 0$, $\varepsilon \in \mathbb{R}$. Set

$$S = \{h \in {}^*\mathbb{R}^+ | h \leqslant m_i \leqslant M, \text{each } i, \Rightarrow |f(m) - a| < \varepsilon\}$$

S is internal.

$(0, M] \setminus F \subset S$ where F is the set of finite numbers. But $(0, M] \setminus F$ is external. So $\exists h, \in S \cap F$. But then

$$m_i > {}^0h_i + 1 \qquad i \in \{1, \dots\} \Rightarrow |f(m) - a| < \varepsilon$$

i.e.

$$\lim_{\substack{m_i \to \infty \\ \text{each } i}} f(m_1, \dots, m_2) = a.$$

We have established that given $\varepsilon > 0$, $\varepsilon \in \mathbb{R}$ there exists $N \in \mathbb{R}$ such that we can write the sentence

$$(\forall m_i \in \mathbb{R})[((\forall i \in \{1, \ldots, n\})(m_i > N)) \Rightarrow |f(m) - a| < \varepsilon].$$

We use the transfer principle to conclude that $\forall \varepsilon > 0$, $\varepsilon \in \mathbb{R}$ $|f(m) - a| < \varepsilon$ whenever m_i are infinite. So $f(m) \approx a$ whenever m_i are infinite.

References

Blanchard Ph and Tarski J 1978 *Acta Phys. Austriaca* **49** 129–52

Bjorken J D and Drell S D 1964 *Relativistic Quantum Mechanics* (New York: McGraw-Hill)

Davis M 1977 *Applied Non-standard Analysis* (New York: Wiley)

Faroukh M O 1975 *J. Math. Phys.* **16** 177

Keleman P and Robinson A 1972 *J. Math. Phys* **13** 1870–8

Luxemburg W A J 1962 (revised 1964) *Non-Standard Analysis: Lectures on Robinson's theory of infinitesimals and Infinitely Large Numbers.* (Pasadena)

—— 1973 *What is Non-Standard Analysis? Papers in the Foundations of Mathematics. Am. Math. Mon.*

Machover M and Hirschfield J 1969 *Lectures On non-Standard Analysis, Lecture notes in Mathematics*, No 94 (Berlin: Springer)

Olver F W J 1974 *Asymptotics and Special Functions* (New York: Academic)

Robinson A 1966 *Non-Standard Analysis: Studies in Logic and the Foundations of Mathematics* (Amsterdam: North-Holland)

Stroyan K D and Luxemburg W A J 1976 *Introduction to the Theory of Infinitesimals* (New York: Academic)

Tarski J 1978 *Many Degrees of Freedom in Field Theory* ed L Streit (New York: Plenum) pp 225–39

Voros A 1974 *J. Math. Phys.* **14** 292–6