Simplified production of

DIRAC DELTA FUNCTION IDENTITIES

Nicholas Wheeler, Reed College Physics Department
November 1997

Introduction. To describe the smooth distribution of (say) a unit mass on the x-axis, we introduce distribution function $\mu(x)$ with the understanding that

$$\mu(x)\,dx \equiv \text{mass element } dm \text{ in the neighborhood } dx \text{ of the point } x$$

$$\int \mu(x)\,dx = 1$$

To describe a mass distribution *localized to the vicinity of* $x = a$ we might, for example, write

$$\mu(x-a;\epsilon) = \begin{cases} \frac{1}{2\epsilon} \text{ if } a-\epsilon < x < a+\epsilon, \text{ and } 0 \text{ otherwise; else} \\ \frac{1}{\sqrt{2\pi\epsilon}}\exp\left\{-\frac{1}{2\epsilon}(x-a)^2\right\}; \qquad \text{else} \\ \frac{1}{\pi x}\sin(x/\epsilon); \qquad \text{else} \ldots \end{cases}$$

In each of those cases we have $\int \mu(x-a;\epsilon)\,dx = 1$ for all $\epsilon > 0$, and in each case it makes formal sense to suppose that

$$\lim_{\epsilon\downarrow 0}\mu(x-a;\epsilon) \text{ describes a unit } \textit{point} \text{ mass situated at } x = a$$

Dirac clearly had precisely such ideas in mind when, in §15 of his *Quantum Mechanics*,[1] he introduced the point-distribution $\delta(x-a)$. He was well aware

[1] I work from his Revised 4th Edition (1967), but the text is unchanged from the 3rd Edition (1947). Dirac's first use of the δ-function occurred in a paper published in 1926, where $\delta(x-y)$ was intended to serve as a continuous analog of the Kronecker delta δ_{mn}, and thus to permit unified discussion of discrete and continuous spectra.

that the "delta function"—which he presumes to satisfy the conditions

$$\int_{-\infty}^{+\infty} \delta(x-a)\,dx = 0$$
$$\delta(x-a) = 0 \quad \text{for} \quad x \neq a$$

—is "not a function. . . according to the usual mathematical definition;" it is, in his terminology, an "improper function," a notational device intended to by-pass distracting circumlocutions, the use of which "must be confined to certain simple types of expression for which it is obvious that no inconsistency can arise."

Dirac's cautionary remarks (and the efficient simplicity of his idea) notwithstanding, some mathematically well-bred people did from the outset take strong exception to the δ-function. In the vanguard of this group was John von Neumann, who dismissed the δ-function as a "fiction," and wrote his monumental *Mathematische Grundlagen der Quantenmechanik*[2] largely to demonstrate that quantum mechanics can (with sufficient effort!) be formulated in such a way as to make no reference to such a fiction.

The situation changed, however, in 1950, when Laurent Schwartz published the first volume of his demanding multi-volume *Théorie des distributions*. Schwartz' accomplishment was to show that δ-functions are (not "functions," either proper or "improper," but) mathematical objects of a fundamentally new type—"distributions," that live always in the shade of an implied integral sign. This was comforting news for the physicists who had by then been contentedly using δ-functions for thirty years. But it was news without major consequence, for Schwartz' work remained inaccessible to all but the most determined of mathematical physicists.

Thus there came into being a tradition of simplification and popularization. In 1949 Schwartz gave a series of lectures at the Seminar of the Canadian Mathematical Congress (held in Vancouver, B.C.), which gave rise in 1952 to a pamphlet[3] that circulated widely, and brought at least the essential elements of the theory of distributions into such clear focus as to serve the simple needs of non-specialists. In 1955 the British applied mathematician G. Temple—building upon remarks published a few years earlier by Mikusínski[4]—published what he called a "less cumbersome vulgarization" of Schwartz' theory, which he hoped might better serve the practical needs of engineers and physicists. Temple's lucid paper inspired M. J. Lighthill to write the monograph from which many of the more recent "introductions to the theory of distributions"

[2] The German edition appeared in 1932. I work from the English translation of 1955. Remarks concerning the δ-function can be found in §3 of Chapter I.

[3] I. Halperin, *Introduction to the Theory of Distributions*.

[4] J. G. Mikusínski, "Sur la méthode de généralization de Laurent Schwartz et sur la convergence faible," Fundamenta Mathematicae **35**, 235 (1948).

descend. Lighthill's slender volume[5]—by intention a text for undergraduates—bears this dedication

TO
PAUL DIRAC
who saw that it must be true
LAURENT SCHWARTZ
who proved it,
AND
GEORGE TEMPLE
who showed how simple it could be made

and has about it—as its title promises—a distinctly "Fourier analytic" flavor. Nor is this fact particularly surprising; the Fourier integral theorem

$$f(x) = \tfrac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ikx}\Big\{\int_{-\infty}^{+\infty} e^{+iky} f(y)\,dy\Big\}\,dk \quad \text{for "nice" functions } f(\bullet)$$

can by reorganization be read as an assertion that

$$\tfrac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ik(y-x)}\,dk = \delta(y-x)$$

The history of the δ-function can in this sense be traced back to the early 1820's. Fourier, of course, was concerned with the theory of heat conduction, but by 1828 the δ-function had intruded for a second time into a physical theory; George Green noticed that the solution of the Poisson equation $\nabla^2\varphi(x) = \rho(x)$, considered to describe the electrostatic potential generated by a given charge distribution $\rho(x)$, can be obtained by superposition of the potentials generated by a population of point charges; i.e., that the general problem can be reduced to the special problem

$$\nabla^2\varphi(x;y) = \delta(x-y)$$

where now the δ-function is being used to describe a "unit point charge positioned at the point y." Thus came into being the "theory of Green's functions," which—with important input by Kirchhoff (physical optics, in the 1880's) and Heaviside (transmission lines, in the 1890's)—became, as it remains, one of the principal consumers of applied distribution theory.

I have sketched this history[6] in order to make clear that what I propose to do in these pages stands quite apart, both in spirit and by intent, from the trend of recent developments, and is fashioned from much ruder fabric. My objective is to promote a point of view—a *computational technique*—that came

[5] *Introduction to Fourier Analysis & Generalized Functions* (1958).

[6] Of which Jesper Lützen, in his absorbing *The Prehistory of the Theory of Distributions* (1982), provides a wonderfully detailed account. In his Concluding Remarks Lützen provides a nicely balanced account of the relative contributions of Schwartz and of S. L. Sobelev (in the early 1930's).

accidentally to my attention in the course of work having to do with the one-dimensional theory of waves.[7] I proceed very informally, and will be concerned not at all with precise characterization of the conditions under which the things I have to say may be true; this fact in itself serves to separate me from recent tradition in the field.

Regarding my specific objectives... Dirac remarks that "There are a number of elementary equations which one can write down about δ functions. These equations are essentially *rules of manipulation* for algebraic work involving δ functions. The meaning of any of these equations is that its two sides give equivalent results [when used] as *factors in an integrand*. Examples of such equations are

$$\begin{aligned}
\delta(-x) &= \delta(x) \\
x\delta(x) &= 0 \\
\delta(ax) &= a^{-1}\delta(x) && : \ a > 0 \qquad (1.1)\\
\delta(x^2 - a^2) &= \tfrac{1}{2}a^{-1}\big\{\delta(x-a) + \delta(x+a)\big\} && : \ a > 0 \qquad (1.2)\\
\textstyle\int \delta(a-x)\,dx\,\delta(x-b) &= \delta(a-b) \\
f(x)\delta(x-a) &= f(a)\delta(x-a) \text{ ''}
\end{aligned}$$

On the evidence of this list (which attains the length quoted only in the 3rd edition) Lützen concludes that "Dirac was a skillful manipulator of the δ-function," and goes on to observe that "some of the above theorems, especially (1.2), are not even obvious in distribution theory, since the changes of variables are hard to perform..."[8] The formal identities in Dirac's list are of several distinct types; he supplies an outline of the supporting argument in all cases but one: concerning (1) he remarks only that they "may be verified by similar elementary arguments." But the elementary argument that makes such easy work of (1.1)[9] acquires a fussy aspect when applied to expressions of the form $\delta\big(g(x)\big)$ typified by the left side of (1.2). My initial objective will be to demonstrate that certain kinds of δ-identities (including particularly those of type (1)) *become trivialities when thought of as corollaries of their θ-analogs*. By extension of the method, I will then derive relationships among the derivative properties of $\delta(\bullet)$ which are important to the theory of Green's functions.

---

[7] See R. Platais, "An investigation of the acoustics of the flute" (Reed College physics thesis, 1993).

[8] See Chapter 4, §29 in the monograph previously cited.

[9] By change of variables we have

$$\begin{aligned}
\int f(x)\delta(ax)\,dx = \int f(y/a)\delta(y)\tfrac{1}{|a|}dy &= \tfrac{1}{|a|}f(0) \\
&= \int f(y)\tfrac{1}{|a|}\delta(y)\,dy
\end{aligned}$$

which assumes only that the Jacobian $|a| \neq 0$.

1. Properties and applications of the Heaviside step function. The step function $\theta(\bullet)$—introduced by Heaviside to model the action of a simple switch—can be defined

$$\theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{2} & \text{at } \ \ x = 0 \\ 1 & \text{for } x > 0 \end{cases} \tag{2}$$

where the central $\frac{1}{2}$ is a (usually inconsequential) formal detail, equivalent to the stipulation that

$$\varepsilon(x) \equiv 2\theta(x) - 1 = \begin{cases} -1 & \text{for } x < 0 \\ \ \ 0 & \text{at } \ \ x = 0 \\ +1 & \text{for } x > 0 \end{cases} \tag{3}$$

be odd (i.e., that $\varepsilon(\bullet)$ vanish at the origin). As Dirac himself (and before him Heaviside) have remarked, the step function (with which Dirac surely became acquainted as a student of electrical engineering) and the δ-function stand in a close relationship supplied by the calculus:

$$\theta(x-a) = \int_{-\infty}^{x} \delta(y-a)\,dy \tag{4.1}$$

$$\Updownarrow$$

$$\tfrac{d}{dx}\theta(x-a) = \delta(x-a) \tag{4.2}$$

The central $\frac{1}{2}$ is, in this light, equivalent to the stipulation that $\delta(x)$ be (formally) an even function of x. For the same reason that $\delta(x)$ becomes meaningful only "in the shade of an integral sign," so also does $\theta(x)$, at least as it is used in intended applications; the construction

$$\delta(x) \equiv \lim_{\epsilon\downarrow 0} \delta(x;\epsilon) \quad \text{entails} \quad \theta(x) = \lim_{\epsilon\downarrow 0} \left\{ \theta(x;\epsilon) \equiv \int_{-\infty}^{x} \delta(y;\epsilon)\,dy \right\}$$

and causes $\theta(x;\epsilon)$ to become literally differentiable at the origin, except in the limit.

That $\delta(x)$ and $\theta(x)$ are complementary constructs can be seen in yet another way. The identity

$$f(x) = \int_{-\infty}^{+\infty} \delta(x-y) f(y)\,dy \tag{5}$$

provides what might be called the "picket fence representation" of $f(x)$. But

$$\begin{aligned} &= \int_{-\infty}^{+\infty} \big\{ -\tfrac{d}{dy}\theta(x-y) \big\} f(y)\,dy \\ &= \int_{-\infty}^{+\infty} \theta(x-y) f'(y)\,dy + \text{boundary term} \end{aligned} \tag{6}$$

which (under conditions that cause the boundary term to vanish) provides the less frequently encountered "stacked slab representation" of $f(x)$. In the former it is $f(\bullet)$ itself that serves to regulate the "heights of successive pickets, while in the latter it is not f but its derivative $f'(\bullet)$ that regulates the "thicknesses of successive slabs." For graphical representations of (5) and (6) see Figure 1.

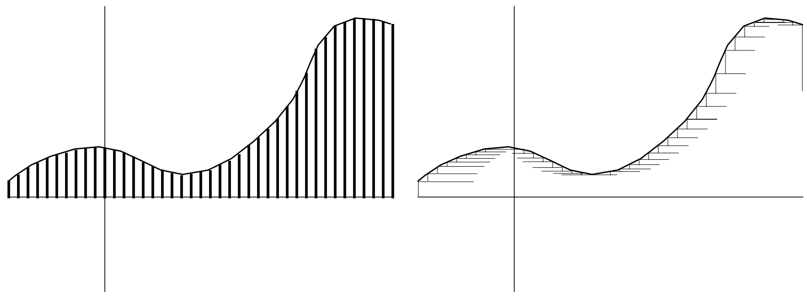

FIGURE 1: *The "picket fence representation" (5) of $f(x)$, compared with the "stacked slab representation" (6).*

Partial integration (subject always to the presumption that boundary terms vanish), which we used to obtain (6), is standardly used also to assign meaning to the successive derivatives of the δ-function; one writes

$$\begin{aligned}
\int f(y)\delta'(y-x)\,dy &= -\int f'(y)\delta(y-x)\,dy = -f'(x)\\
\int f(y)\delta''(y-x)\,dy &= -\int f'(y)\delta'(y-x)\,dy\\
&= (-)^2\int f''(y)\delta(y-x)\,dy = (-)^2 f''(x)\\
&\vdots\\
\int f(y)\delta^{(n)}(y-x)\,dy &= (-)^n f^{(n)}(x)
\end{aligned}$$

Formally, one has (see Figure 2)

$$\left.\begin{aligned}
\delta'(y-x) &= \lim_{\epsilon\downarrow 0}\frac{\delta\big(y-(x-\epsilon)\big)-\delta\big(y-(x+\epsilon)\big)}{2\epsilon}\\
\delta''(y-x) &= \lim_{\epsilon\downarrow 0}\frac{\delta\big(y-(x-2\epsilon)\big)-2\delta\big(y-x\big)+\delta\big(y-(x+2\epsilon)\big)\big)}{(2\epsilon)^2}\\
&\ \ \vdots
\end{aligned}\right\}\tag{7}$$

which when introduced as factors into the integrand of $\int f(y)\,dy$ serve simply to reconstruct the *definitions* of $f'(x)$, $f''(x)$, etc.

In anticipation of future needs, I note that the preceding equations can be formulated as relations among θ-functions. Taking (4.2) as our point of departure, we have

$$\begin{aligned}
\delta(y-x) &= \lim_{\epsilon\downarrow 0}\frac{\theta\big(y-(x-\epsilon)\big)-\theta\big(y-(x+\epsilon)\big)}{2\epsilon}\\
&= \text{limit of a sequence of ever-narrower/taller "box functions"}
\end{aligned}$$

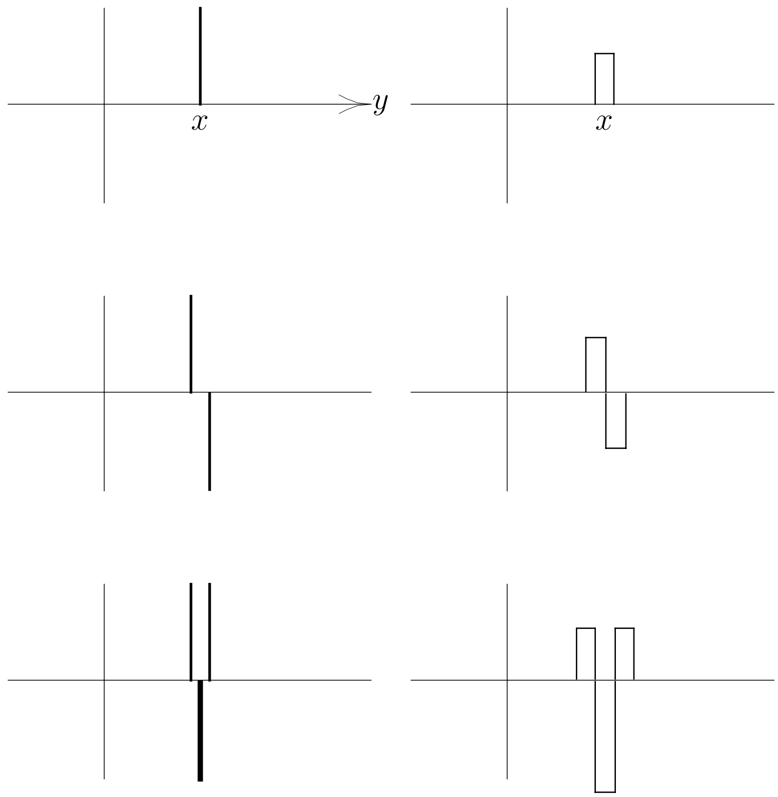

FIGURE 2: *The figures on the left derive from (7), and show δ representations of ascending derivatives of $\delta(y-x)$. The figures on the right derive from (8), and provide θ representations of the same material.*

giving (see again the preceding figure)

$$\left.\begin{aligned}
\delta'(y-x) &= \lim_{\epsilon\downarrow 0}\tfrac{\theta(y-[x-2\epsilon])-2\theta(y-x)+\delta(y-[x+2\epsilon])}{(2\epsilon)^2}\\
\delta''(y-x) &= \lim_{\epsilon\downarrow 0}\tfrac{\theta(y-[x-3\epsilon])-3\theta(y-[x-\epsilon])+3\delta(y-[x+\epsilon])-\theta(y-[x+3\epsilon])}{(2\epsilon)^3}\\
&\vdots
\end{aligned}\right\}\tag{8}$$

2. Simplified derivation of delta function identities. Let $\theta(x;\epsilon)$ refer to some (any nice) parameterized sequence of functions convergent to $\theta(x)$, and let a be a positive constant. While $\theta(ax;\epsilon)$ and $\theta(x;\epsilon)$ are distinct functions of x, they clearly become *identical in the limit* $\epsilon\downarrow 0$, and so also therefore do their derivatives (of all orders). So we have $a\theta'(ax)=\theta'(x)$, which by (4.2) reproduces (1.1):

$$\delta(ax)=a^{-1}\delta(x)\quad : \quad a>0\tag{9.1}$$

If, on the other hand, $a < 0$ then $\theta(ax) = 1 - \theta(x)$ gives $a\theta'(ax) = -\theta'(x)$ whence

$$\delta(ax) = -a^{-1}\delta(x) \quad : \quad a < 0 \tag{9.2}$$

The equation

$$\delta(ax) = \tfrac{1}{|a|}\,\delta(x) \quad : \quad a \neq 0 \tag{10.1}$$

provides a unified formulation of (9.1) and (9.2). A second differentiation gives

$$\begin{aligned} a^2\theta''(ax) &= \pm\theta''(x) \quad \text{according as } a \gtrless 0 \\ &\Downarrow \\ \delta'(ax) &= \pm\tfrac{1}{a^2}\,\delta'(x) \end{aligned} \tag{10.2}$$

Preceding remarks illustrate the sense in which "δ-identities become trivialities when thought of as corollaries of their θ-analogs." By way of more interesting illustration of the same point...

The function $g(x) \equiv x^2 - a^2$ describes an up-turned parabola that crosses the x-axis at $x = \pm a$. Evidently

$$\begin{aligned} \theta(x^2 - a^2) &= \begin{cases} 1 & \text{for } x < -a \\ 0 & \text{for } -a < x < +a \\ 1 & \text{for } x > +a \end{cases} \\ &= 1 - \big\{\theta(x+a) - \theta(x-a)\big\} \end{aligned}$$

by differentiation entails

$$\begin{aligned} 2x\,\delta(x^2 - a^2) &= -\delta(x+a) + \delta(x-a) \\ &\downarrow \\ \delta(x^2 - a^2) &= -(-2a)^{-1}\delta(x+a) + (+2a)^{-1}\delta(x-a) \\ &= \tfrac{1}{2|a|}\Big\{\delta(x-a) + \delta(x+a)\Big\} \end{aligned} \tag{11}$$

which provides a generalized formulation if (1.2). By this mode of argument it becomes transparently clear how the a that enters into the prefactor comes to acquire its (otherwise perplexing) absolute value braces.

Suppose, more generally, that

$$g(x) = g_0(x - x_1)(x - x_2)\cdots(x - x_n)$$

with $x_1 < x_2 < \cdots < x_n$. We then (see Figure 3) have

$$\theta(g(x)) = \begin{cases} \theta(x - x_1) - \theta(x - x_2) + \cdots - (-)^n\theta(x - x_n) \\ \\ 1 - \big\{\theta(x - x_1) - \theta(x - x_2) + \cdots - (-)^n\theta(x - x_n)\big\} \end{cases} \tag{12}$$

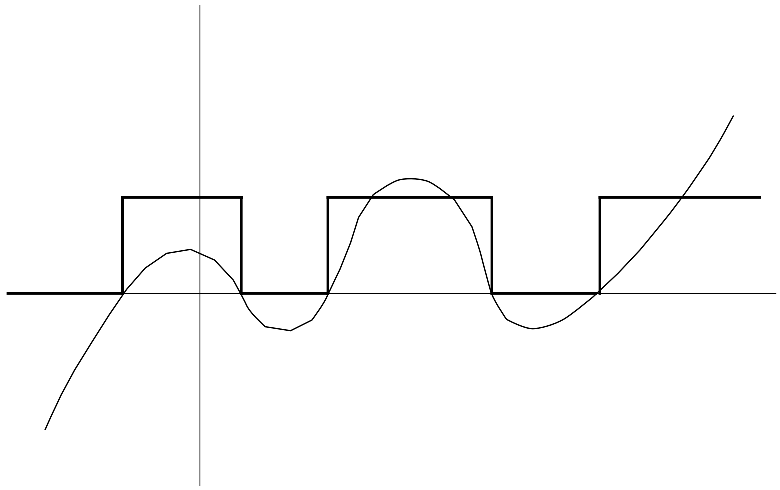

FIGURE 3: *Representation of $\theta(g(x))$ in a typical case.*

according as $g'(x_1) \gtrless 0$. By differentiation

$$\begin{aligned} g'(x)\cdot\delta(g(x)) &= \pm\Big\{\delta(x-x_1)-\delta(x-x_2)+\cdots-(-)^n\delta(x-x_n)\Big\}\\ &\downarrow\\ \delta(g(x)) &= \sum_{k=1}^{n}\tfrac{1}{|g'(x_k)|}\delta(x-x_k)\quad\text{in all cases} \end{aligned}\tag{13}$$

The identities (10.1) and (11) are seen to comprise special instances of this famous result. The more standard derivation involves massaging the integral $\int f(x)\delta(g(x))\,dx$, and is a good deal less immediate. The approach here advocated owes its simplicity very largely to the simplicity of (12).

3. Elementary geometrical approach to properties of some Green's functions. It was known already to d'Alembert (1747) that if $\varphi(t,x)$ is a solution of the wave equation[10]

$$\Box\varphi = 0 \quad\text{with}\quad \Box\equiv\partial_t^2-\partial_x^2$$

and if initial data $\varphi(0,x)$ and $\varphi_t(0,x)$ are prescribed, then

$$\begin{aligned} \varphi(t,x) &= \tfrac{1}{2}\Big\{\varphi(0,x-t)+\varphi(0,x+t)\Big\}+\tfrac{1}{2}\int_{x-t}^{x+t}\varphi_t(0,y)\,dy \\ &= \text{function of }(x-t)+\text{function of }(x+t)\\ &= \text{right-running wave}+\text{left-running wave} \end{aligned}\tag{14}$$

[10] For the purposes of this discussion I find it convenient (though alien to my practice) to set $c=1$.

But two centuries were to elapse before it came to be widely appreciated that (14) can usefully be notated

$$\varphi(t,x)=\int_{-\infty}^{+\infty}\Big\{\nabla_0(x-y,t)\varphi(0,y)+\Delta^0(x-y,t)\varphi_t(0,y)\Big\}\,dy \tag{15}$$

with

$$\Delta^0(x-y,t)\equiv\tfrac{1}{2}\Big\{\theta\big(y-(x-t)\big)-\theta\big(y-(x+t)\big)\Big\} \tag{16.1}$$

$$\begin{aligned}\nabla_0(x-y,t)&\equiv\tfrac{1}{2}\Big\{\delta\big(y-(x-t)\big)+\delta\big(y-(x+t)\big)\Big\} \qquad (16.2)\\ &=\tfrac{\partial}{\partial t}\,\Delta^0(x-y,t)\end{aligned}$$

It is important to notice that, while we had *prediction* (i.e., evolution from prescribed *initial* data) in mind when we devised (15), the equation also works *retrodictively*; (15) is invariant under $t\longmapsto -t$. Under time-reversal all ∂_t's change sign, but so also (manifestly) does $\Delta^0(x-y,t)$, and when we return with this information to (15) we see that the two effects precisely cancel. Evidently $\Delta^0(x-y,t-u)$—thought of as a function of $\{u,y\}$ that depends parametrically on $\{t,x\}$—has the properties

$$\begin{aligned}\Delta^0(x-y,t-u)&=\begin{cases}+\tfrac{1}{2}\text{ inside the lightcone extending backward from }\{t,x\}\\ -\tfrac{1}{2}\text{ inside the lightcone extending forward from }\{t,x\}\\ \;\;0\text{ outside the lightcone}\end{cases}\\ &=\tfrac{1}{2}\Big\{\theta\big((y-x)+(t-u)\big)-\theta\big((y-x)-(t-u)\big)\Big\}\\ &=\tfrac{1}{2}\varepsilon(t-u)\cdot\,\theta\big((t-u)^2-(x-y)^2\big)\end{aligned}$$

illustrated in Figure 4. In order for the right side of (15) to do its assigned work it must be the case that

$$\big\{\partial_t^2-\partial_x^2\big\}\Delta^0(x-y,t-u)=0$$

and

$$\begin{aligned}\Delta^0\;(x-y,0)&=0\\ \Delta^0_t\;(x-y,0)&=\delta(y-x)\\ \Delta^0_{tt}(x-y,0)&=0\end{aligned}$$

—all of which are, in fact, quick consequences of (16).

The *forced* wave equation reads

$$\Box\varphi=F\quad\text{with }F(t,x)\text{ prescribed}$$

The general solution can be developed

$$\varphi=\varphi_0+\varphi_F$$

where φ_0 evolves according to $\Box\varphi_0=0$ from prescribed initial data, while φ_F evolves according to $\Box\varphi_F=F$ but is assumed to vanish on the initial timeslice.

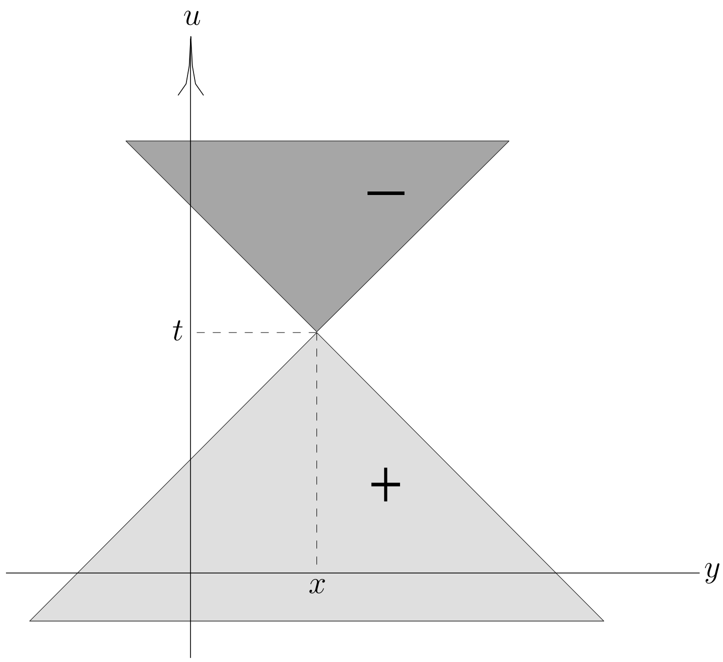

FIGURE 4: *Representation of the Green's function* $\Delta^0(x-y,t-u)$ *of the homogeneous wave equation* $\Box\varphi = 0$*. For* $u < t$ *the function has the form of a triangular plateau (backward lightcone) with a flat top at elevation* $\frac{1}{2}$*, while for* $u > t$ *(forward lightcone) it is a triangular excavation of similar design.*

Green's method leads one to write

$$\varphi_{\scriptscriptstyle F}(t,x) = \iint \Delta(x-y,t-u)F(u,y)\,dydu$$

and to require of $\Delta(x-y,t-u)$ that

$$\big\{\partial_t^2 - \partial_x^2\big\}\Delta(x-y,t-u) = \delta(y-x)\delta(u-t) \tag{17}$$

and $\Delta(y-x,0) = 0$. Such properties are possessed by (in particular) this close relative of Δ^0

$$\begin{aligned}
\Delta(x-y,t-u) &= \theta(t-u)\cdot\Delta^0(x-y,t-u) \qquad (18)\\
&= \begin{cases} \frac{1}{2}\big[\theta\big((y-x)+(t-u)\big) - \theta\big((y-x)-(t-u)\big)\big] & : \quad t-u>0 \\ 0 & : \quad t-u<0 \end{cases}\\
&= \begin{cases} +\frac{1}{2} \text{ inside the lightcone extending backward from } \{t,x\} \\ 0 \text{ inside the lightcone extending forward from } \{t,x\} \\ 0 \text{ outside the lightcone} \end{cases}
\end{aligned}$$

which is known as the "retarded propagator." The properties in question are possessed also by the similarly-defined "advanced propagator," and are therefore shared also by *all* functions of the form[11]

$$\Delta(\bullet,\bullet) = k\Delta_{\text{ret}}(\bullet,\bullet) + (1-k)\Delta_{\text{adv}}(\bullet,\bullet)$$

We encounter at this point the need for a "principle of choice," but it would distract me from my primary objective to pursue the matter.

Concerning that "primary objective:" I entered into the preceding review of what is after all standard material partly to demonstrate that the Fourier analytic methods standardly encountered[12] are, in fact, inessential (the subject is susceptible to development by elementary means)... but mainly to place

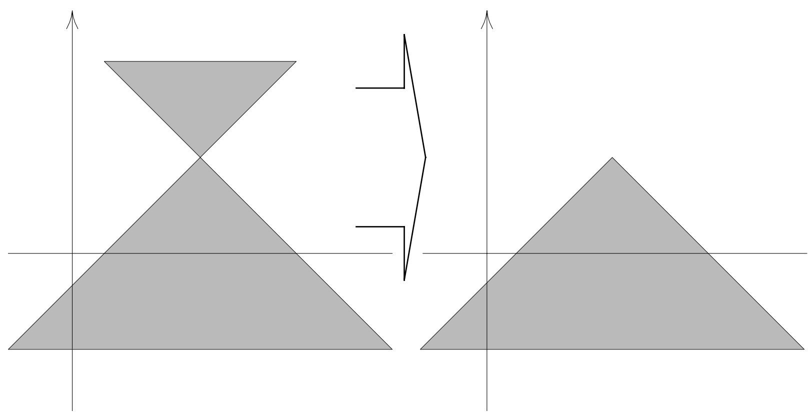

FIGURE 5: *Representation of the "truncation" process (18), by means of which Δ^0 gives rise to Δ.*

myself in position to pose this question: *What is the detailed mechanism by means of which "truncation" causes a solution Δ^0 of the homogeneous wave equation to become a solution Δ of the inhomogeneous equation:*

$$\Box\Delta^0 = 0 \quad \xrightarrow[\text{truncation of }\Delta^0]{} \quad \Box\Delta = \delta$$

Elementary insight into the geometrical essence of the answer is provided by the following sequence of figures:

[11] See F. Rohrlich, *Classical Charged Particles* (1965), p. 79.

[12] See, for example, RELATIVISTIC CLASSICAL FIELDS (1973), pp. 163–190 and ANALYTICAL METHODS OF PHYSICS (1981), pp. 291–297.

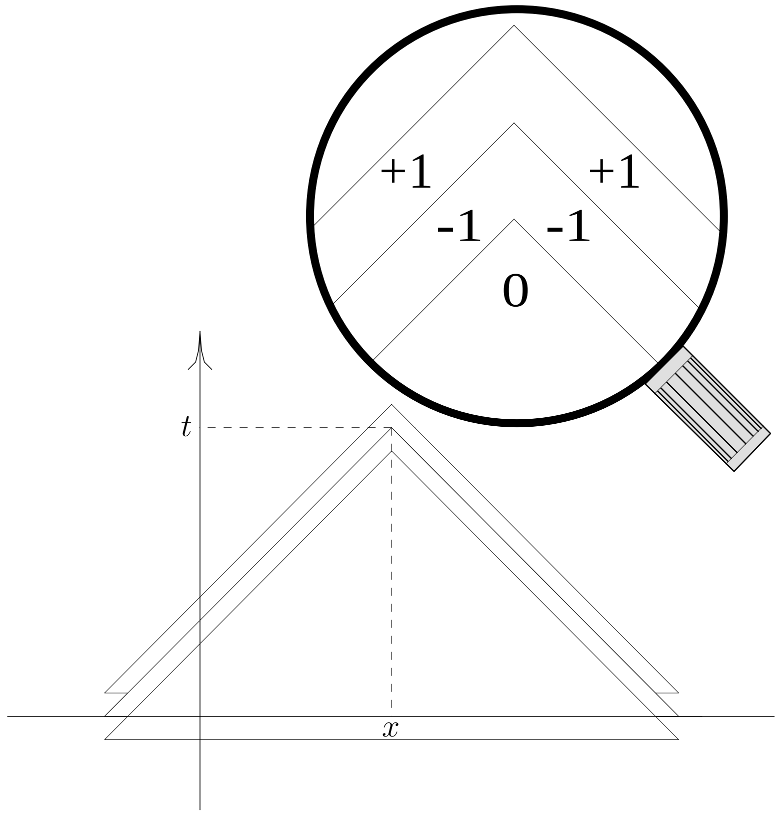

FIGURE 6: *The numbers, when multiplied by* $\frac{1}{2}(2\epsilon)^{-2}$, *refer—in the sense familiar from (7) and (8); see also the lower right detail in Figure 2—to the discrete approximation*

$$+\partial_t^2\Delta(x-y,t-u) \approx +\frac{\Delta(x-y,[t+2\epsilon]-u) - 2\Delta(x-y,[t]-u) + \Delta(x-y,[t-2\epsilon]-u)}{(2\epsilon)^2}$$

In this and subsequent figures $\Delta(x-y,t-u)$ *is considered to be a function of* $\{u,y\}$, *into which* $\{t,x\}$ *enter as parameters.*

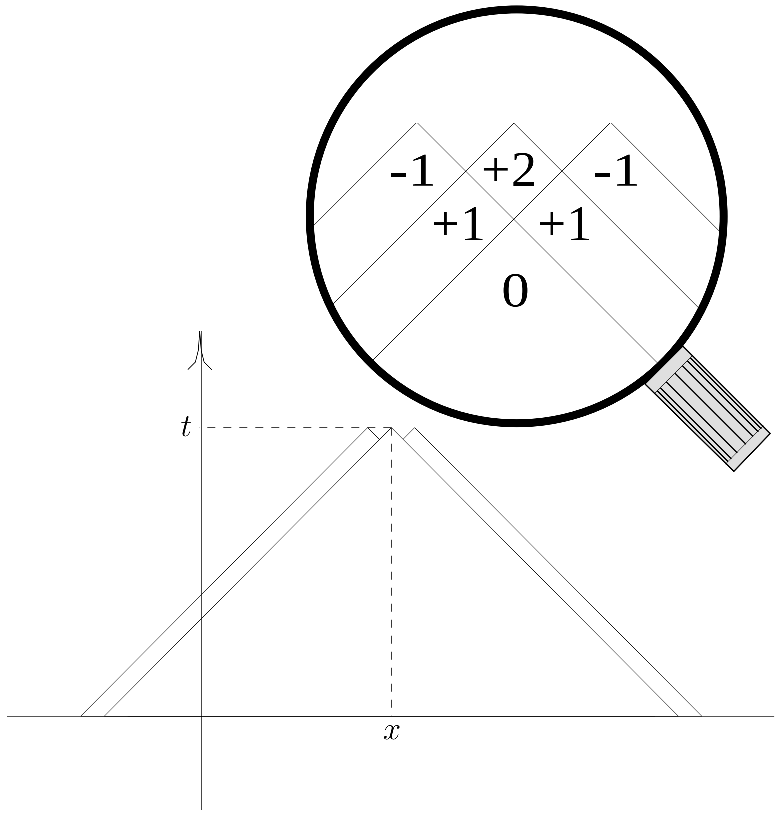

FIGURE 7: *The numbers, when multiplied by* $\frac{1}{2}(2\epsilon)^{-2}$, *refer to the discrete approximation*

$$-\partial_x^2\Delta(x-y,t-u) \approx -\frac{\Delta([x+2\epsilon]-y,t-u)-2\Delta([x]-y,t-u)+\Delta([x-2\epsilon]-y,t-u)}{(2\epsilon)^2}$$

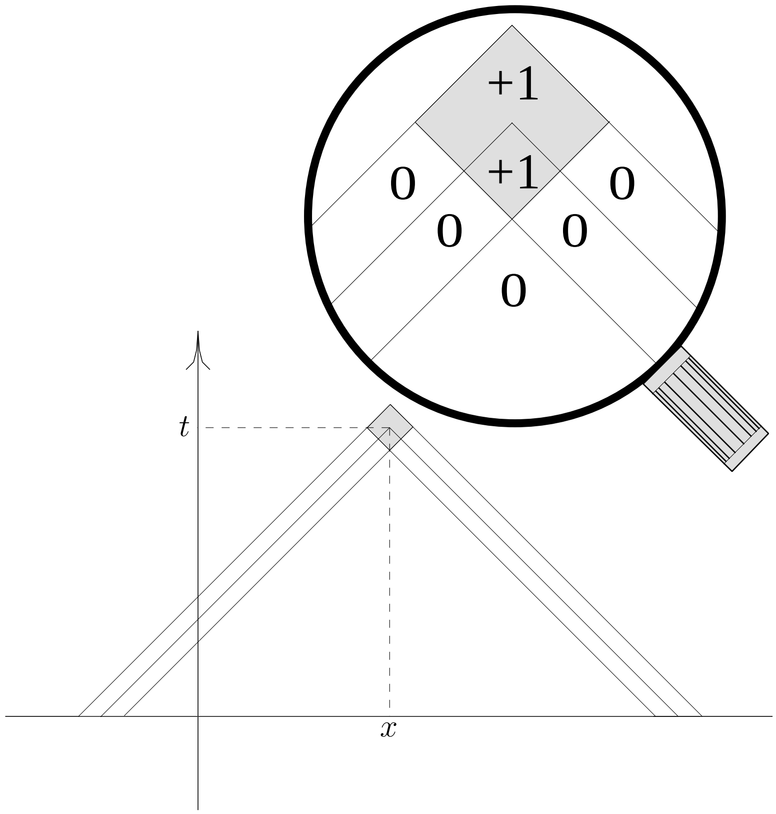

FIGURE 8: *Superposition of the preceding figures. The numbers, when multiplied by* $\frac{1}{2}(2\epsilon)^{-2}$*, supply the discrete approximation*

$$\left\{\partial_t^2 - \partial_x^2\right\}\Delta(x-y, t-u) \approx \tfrac{1}{2}(2\epsilon)^{-2}\cdot \mathrm{Box}(t,x;\epsilon)$$

where

$$\mathrm{Box}(t,x;\epsilon) \equiv \begin{cases} 1 \text{ interior to shaded box centered at } (t,x) \\ 0 \text{ elsewhere} \end{cases}$$

The box has diagonal measure 4ϵ*, and its area is given therefore by* $(4\epsilon/\sqrt{2})^2 = 8\epsilon^2$*. It is the top of a little prism of height* $\frac{1}{2}(2\epsilon)^{-2}$*, the volume of which is given by*

$$\text{volume} = 1 \quad \text{(all values of } \epsilon)$$

evidently

$$\lim_{\epsilon\downarrow 0} \tfrac{1}{2}(2\epsilon)^{-2}\cdot \mathrm{Box}(t,x;\epsilon) = \delta(u-t)\delta(y-x)$$

which yields the result we sought to establish.

4. Analytical derivation of the same result. The discussion in the preceding section culminated in a geometrical argument designed to illuminate how it comes about that

$$\Box\Delta^0 = 0 \quad \text{but truncation engenders} \quad \Box\Delta = \delta$$

For comparative purposes I now outline the analytical demonstration of that same fact; the details are not without interest, but the argument as a whole seems to me to lack the "ah-ha!" quality of its geometrical counterpart.

By (18) we have

$$\begin{aligned}\Delta(x-y,t-u) &= \tfrac{1}{2}\theta(t-u)\theta(\&c.)\\ &\qquad\qquad \&c. \equiv (t-u)^2-(x-y)^2\end{aligned}$$

Therefore

$$\begin{aligned}\partial_t\Delta &= \tfrac{1}{2}\delta(t-u)\theta(\&c.) \ +\theta(t-u)(t-u)\delta(\&c.)\\ \partial_t\partial_t\Delta &= \tfrac{1}{2}\delta'(t-u)\theta(\&c.)+2\delta(t-u)(t-u)\delta(\&c.)\\ &\qquad +\theta(t-u)\delta(\&c.)+\theta(t-u)2(t-u)^2\delta'(\&c.)\\ \partial_x\Delta &= \qquad\qquad -\theta(t-u)(x-y)\delta(\&c.)\\ \partial_x\partial_x\Delta &= \qquad\qquad -\theta(t-u)\delta(\&c.)+\theta(t-u)2(x-y)^2\delta'(\&c.)\end{aligned}$$

giving

$$\begin{aligned}\big\{\partial_t^2-\partial_x^2\big\}\Delta &= \tfrac{1}{2}\delta'(t-u)\theta(\&c.)+2\delta(t-u)(t-u)\delta(\&c.)\\ &\qquad +2\theta(t-u)\underbrace{\big\{\delta(\&c.)+(\&c.)\delta'(\&c.)\big\}}_{0}\end{aligned}$$

where the term that vanishes does so because $x\delta(x)=0\Longrightarrow\delta(x)+x\delta'(x)=0$. We note in passing that had we omitted the θ-factor from the definition of Δ; i.e., if we were evaluating $\Box\Delta^0$ instead of $\Box\Delta$, we would at this point have achieved $\Box\Delta^0=0$.

Our assignment now is to establish that

$$\tfrac{1}{2}\delta'(t-u)\theta(\&c.)+2\delta(t-u)(t-u)\delta(\&c.)=\delta(t-u)\delta(x-y)$$

which, if we write $\tau\equiv t-u$ and $\xi\equiv x-y$, can be notated

$$\underbrace{\tfrac{1}{2}\delta'(\tau)\theta(\tau^2-\xi^2)\ +2\delta(\tau)\tau\delta(\tau^2-\xi^2)}=\delta(\tau)\delta(\xi)$$

But

$$=\tfrac{1}{2}\underbrace{\tfrac{\partial}{\partial\tau}\big\{\delta(\tau)\theta(\tau^2-\xi^2)\big\}}+\delta(\tau)\tau\delta(\tau^2-\xi^2)$$

I will argue that

$$=0 \tag{19}$$

Assuming, for the moment, the truth of that claim, we want to show that

$$\underbrace{\delta(\tau)\tau\cdot\delta(\tau^2-\xi^2)}=\delta(\tau)\delta(\xi) \tag{$*$}$$

But this is in fact immediate; looking to (11) and taking advantage of simplifications made available by the presence of the $\delta(\tau)$-factor, we have

$$\begin{aligned}&=\delta(\tau)\tau\cdot\tfrac{1}{2\tau}\big\{\delta(\xi-\tau)+\delta(\xi+\tau)\big\}\\&=\delta(\tau)\cdot\tfrac{1}{2}\big\{\delta(\xi)+\delta(\xi)\big\}\\&=\delta(\tau)\cdot\delta(\xi)\end{aligned}$$

Returning now to the demonstration of (19); if $F(\tau)$ is nice function, then (integrating by parts) we have

$$\begin{aligned}\int\tfrac{\partial}{\partial\tau}\big\{\delta(\tau)\theta(\tau^2-\xi^2)\big\}F(\tau)\,d\tau&=-\int\delta(\tau)\theta(\tau^2-\xi^2)F'(\tau)\,d\tau\\&=-\int\delta(\tau)\big[1-\big\{\theta(\tau+\xi)-\theta(\tau-\xi)\big\}\big]F'(\tau)\,d\tau\quad\text{by (12)}\\&=-\int\delta(\tau)F'(\tau)\,d\tau+\int_{-\xi}^{+\xi}\delta(\tau)F'(\tau)\,d\tau\\&=0\quad\text{all }\xi>0,\ \therefore\text{ also in the limit }\xi\downarrow 0\text{ because }F(\tau)\text{ is “nice”}\end{aligned}$$

The argument just concluded is notable for its delicacy[13] and its overall improvisatory, *ad hoc* quality. It is, in my view, not only less illuminating but also less convincing than the geometrical argument summarized in Figures 6–8. It points up the need for a systematic “calculus of distributions,” and lends force to Dirac’s observation that we should (in the absence of such a calculus) avail ourselves of δ-methods only when “it is obvious that no inconsistency can arise.” This, however, is more easily said than done.

5. Relaxation of the assumption that spacetime is 2-dimensional. In §3 we were able to obtain Green’s functions of the homogeneous/inhomogeneous wave equations

$$\begin{aligned}\big\{\partial_t^2-\partial_x^2\big\}\Delta^0&=0\\\big\{\partial_t^2-\partial_x^2\big\}\Delta\ &=\delta\end{aligned}$$

by putting d’Alembert and Heaviside/Dirac in a pot and stirring gently, the notable fact being that no Fourier was called for by the recipe. d’Alembert made critical use, however, of circumstances which are special to the 2-dimensional

[13] We would, for example, have gone off-track if we had made too-casual use of Dirac’s identity $x\delta(x)=0$; this makes good sense in the context Dirac intended ($\int x\delta(x)\cdot f(x)\,dx=0$), but can lead to error when (as at $(*)$) $x\delta(x)$ appears as a factor in a more complex expression.

case; in higher-dimensional cases d'Alembert's line of argument breaks down, and new analytical methods are called for. Those methods—or at least the methods most commonly brought to bear upon the analysis of

$$\begin{aligned}\big\{\partial_t^2 - \partial_{x^1}^2 - \partial_{x^2}^2 - \cdots - \partial_{x^n}^2\big\}\Delta_n^0(x) &= 0\\ \big\{\partial_t^2 - \partial_{x^1}^2 - \partial_{x^2}^2 - \cdots - \partial_{x^n}^2\big\}\Delta_n(x) &= \delta(x)\end{aligned}$$

—make essential use of Fourier transform and contour integral techniques. The small point to which I wish here to draw attention is that it is nevertheless *possible* to omit Fourier from the recipe in the *general* case; the simple case $n = 1$ provides the "seed" from which results appropriate to the general case can be obtained by elementary operations borrowed from the so-called "fractional calculus." Of course, any person with practical need of (say) $\Delta_n^0(x)$ is a person certain to be comfortable with Fourier analytic methods, grateful for the special insights which those methods typically provide. . . and (in all likelihood) ignorant of the concept of "semi-differentiation." Such a person may be inclined to dismiss the following remarks as an amusing *tour de force*. To do so would, however, be to incur the wrath of the ghosts of Jacques Hadamard and Marcel Riesz, who originally undertook this work not to be cute but to clarify certain deep problems.

One proceeds from the observation[14] that the variables $\{t, x^1, x^2, \ldots, x^n\}$ enter into the structure of $\Delta_n^0(x)$ only in lumped combination

$$\sigma = t^2 - r^2 \quad \text{with} \quad r^2 \equiv x_1^2 + x_2^2 + \cdots + x_n^2$$

and by fairly straightforward argument is led at length to the scheme

$$\begin{array}{ccccccccccc}
\Delta_1 & & \longrightarrow & & \Delta_3 & & \longrightarrow & & \Delta_5 & \longrightarrow & \Delta_7 \\
 & \searrow & & \nearrow & & \searrow & & \nearrow & \quad\searrow & \nearrow & \qquad \ldots\text{etc.}\\
 & & \Delta_2 & & \longrightarrow & & \Delta_4 & & \longrightarrow & \Delta_6 &
\end{array}$$

where horizontal arrows refer to the action of $\frac{1}{\pi}\frac{\partial}{\partial\sigma} = -\frac{1}{2\pi r}\frac{\partial}{\partial r}$ and sloping arrows refer to the action of the semi-differentiation operator $\big(\frac{1}{\pi}\frac{\partial}{\partial\sigma}\big)^{\frac{1}{2}}$. The figure that results when all arrows are reversed makes complementary good sense, provided the reversed arrows are understood to refer to (semi)integration processes. The scheme is computationally very efficient, and it exposes patterns that tend to remain concealed when one proceeds by Fourier analytic means. In particular, it makes evident (when one looks to the details) the profound sense in which physics in odd-dimensional space is to be distinguished from physics in even-dimensional space. And the sense—visible only when one is in position to make comparative statements—in which the case $n = 3$ is "special." We have encountered a situation in which (arguably) "simpler is also deeper." And (which is more immediately germane to the business at hand) we have encountered a context within which not only θ and δ but also δ', δ''. . . enter directly and essentially into the description of the physics.

[14] See §7 of my "Construction & physical application of the fractional calculus" (1997) for mathematical details and physical commentary.