1

ANALYTICAL DYNAMICS OF FIELDS

Introduction. Somewhat idiosyncratically, I like to distinguish the "classical mechanics" of particles from what (until about the time of the appearance—in 1904—of E. T. Whittaker's TREATISE ON THE ANALYTICAL DYNAMICS OF PARTICLES AND RIGID BODIES) used to be called "analytical mechanics." I take the former to embrace *all that can be said, by whatever formal means*, about the dynamics of particulate systems of all descriptions, but understand the latter terminology to refer specifically to the resources latent in the Lagrangian formalism, in the Hamiltonian formalism, in the Hamilton-Jacobi formalism, in their less-well-known companion formalisms (such, for example, as the Appell formalism) and their associated variational principles. Since it is easy to think of systems—particularly, but not exclusively, *dissipative* systems—to which none of the formalisms just ennumerated usefully pertain, it is clear that "analytical mechanics" is by nature a sub-division of a broader discipline. Remarkably, it is (or appears to be) within the confines of that sub-discipline that God prefers to frame His most fundamental utterances.

When one turns from the mechanics of spatially localized systems to the mechanics of *distributed* systems—i.e., from the dynamics of particles to the dynamics of fields—one encounters a similar situation. To think generally of "field theory" is to think of a subject so broad as to embrace all aspects of the motion of fluids, of elastic solids, of electromagnetic and gravitational fields, all—surprisingly—of the quantum mechanics of particles (at least formally), and of much else besides. Within that broad field lives the subject to which my chapter title refers.

Our strategy will be to look to a graded sequence of particulate systems, the limiting member of which will, by design, have in fact the character of a

field. By tracking the analytical mechanics of the individual members of the sequence we shall obtain the analytical mechanics of its limiting member—the "analytical mechanics of a field." The results thus achieved will be so strikingly simple as to admit readily and unproblematically of generalization. This mode of proceeding will, by the way, yield a general-purpose field-theoretic language which is automatically consonant with the analytical mechanics of particles, and is therefore preadapted to the discussion of particles and fields *in interaction*.

Dynamics of one-dimensional crystals. Take N identical particles of mass m, and $N+1$ identical springs of strength k, and form the "crystal" illustrated in the first of the following figures. The terminal springs are attached to "walls" which stand a distance ℓ from one another, so when the system is at rest each

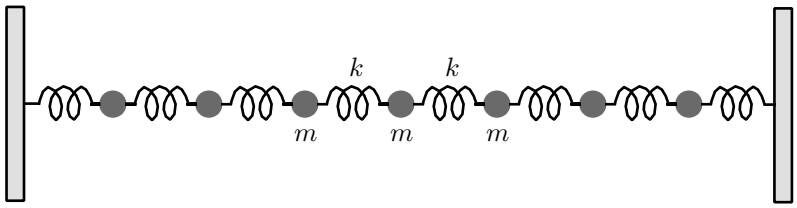


FIGURE 1: *Equilibrium configuration of an N-particle crystal with clamped boundaries.*

spring has length $a = \ell/(N+1)$. The system derives its "one-dimensionality" not so much from the linearity of its design as from the explicit stipulation that *transverse motion will be disallowed*. The allowed motion is longitudinal, which by natural orientation of a Cartesian coordinate system means "along the x-axis." If we associate the origin of the x-axis with the anchor point on the left, then we can write $x_n = na$ to describe the location of the n^{th} "atom" in a crystal at rest. Our dynamical assignment is to develop the functions $x_n(t)$ $(n = 1, 2, \ldots, N)$ which describe the successive locations of the constituent atoms in a crystal *not* at rest.

To that end, let us (see the following figure) introduce "excursion variables" φ_n by means of the equations $x_n = na + \varphi_n$. Evidently φ_n serves to describe

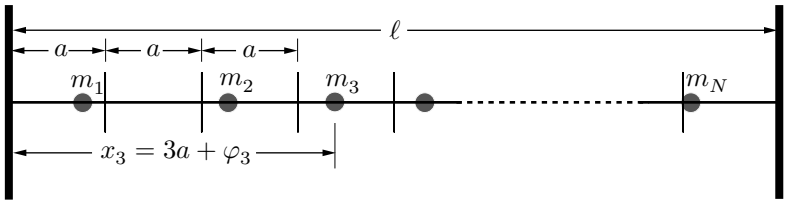


FIGURE 2: *Use of excursion variables to describe the instantaneous configuration of a crystal not at rest.*

the location of the n^{th} atom *relative to its equilibrium position*. It is in precisely the spirit of the familiar "theory of small oscillations" that, in our effort to

comprehend the motion of the crystal, we agree to look to the time-dependence of the variables φ_n.

We invested energy in the formation of our crystal; the interconnecting springs are, after all, stretched (else compressed) in the general case. Energy over and above that "ground state energy" must be invested if we wish to set the crystal in motion. That energy can evidently be described $E = T + U$ with

$$T = \tfrac{1}{2} m \left\{ \dot{\varphi}_1^2 + \dot{\varphi}_2^2 + \cdots + \dot{\varphi}_{N-1}^2 + \dot{\varphi}_N^2 \right\} \tag{1}$$

$$U = \tfrac{1}{2}\, k \, \left\{ \varphi_1^2 + (\varphi_2 - \varphi_1)^2 + \cdots + (\varphi_N - \varphi_{N-1})^2 + \varphi_N^2 \right\} \tag{2}$$

Forming the Lagrangian $L = T - U$ and working from

$$\left\{ \frac{d}{dt} \frac{\partial}{\partial \dot{\varphi}_n} - \frac{\partial}{\partial \varphi_n} \right\} L = 0 \qquad\qquad n = 1, 2, \ldots, N \tag{3}$$

we readily obtain the following explicit equations of motion:

$$\left.\begin{aligned}
m\ddot{\varphi}_1 &= -k(\qquad\quad + 2\varphi_1 - \varphi_2 \quad) \\
m\ddot{\varphi}_2 &= -k(-\varphi_1 \quad + 2\varphi_2 - \varphi_3 \quad) \\
&\vdots \\
m\ddot{\varphi}_n &= -k(-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1}) \qquad n = 2, 3, \ldots, N-1 \\
&\vdots \\
m\ddot{\varphi}_N &= -k(-\varphi_{N-1} + 2\varphi_N \qquad)
\end{aligned}\right\} \tag{4}$$

These comprise a coupled system of N 2^{nd}-order ordinary differential equations in N variables. Equations (4) are notable in particular for their *linearity*, which we might emphasize by writing

$$\ddot{\boldsymbol{\varphi}} + \mathbb{S}\boldsymbol{\varphi} = \mathbf{0} \tag{5}$$

where

$$\boldsymbol{\varphi} \equiv \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \vdots \\ \varphi_N \end{pmatrix} \text{ and } \mathbb{S} \text{ has the structure } \mathbb{S} \equiv \Omega^2 \begin{pmatrix} a & b & & & & & \\ b & a & b & & & & \\ & b & a & b & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & b & a & b & \\ & & & & b & a & \end{pmatrix}$$

with $\Omega^2 = k/m$, $a = 2$ and $b = -1$. The design of the $\mathbb{S}$ matrix (in which all unreported matrix elements are zero) pretty clearly and directly mimics the physical design of the crystal itself.

We have now before us a physical system with a long and important history, an interesting brief account of which can be found in the opening sections of

L. Brillouin's classic *Wave Propagation in Periodic Structures* (1946). The first chapter in that history was written by Newton himself, who used a one-dimensional crystal to model an air column in his pioneering attempt to compute the velocity of sound. The system continues to serve as a point of entry into the study of real crystals (solid state physics), of the vibration of molecules, of transmission lines and of much else. And the system gives rise to an analytical problem of rich methodological interest. It is to aspects of the latter that I am motivated now to give brief attention.

Suppose we return to (5) with the *assumption* that the atoms oscillate with possibly distinct amplitudes A_n but in perfect synchrony:

$$\boldsymbol{\varphi}(t) = \boldsymbol{A}\, e^{i\omega t}$$

Immediately

$$(\mathbb{S} - \omega^2 \mathbb{I}\,)\boldsymbol{A} = \boldsymbol{0}$$

Evidently ω^2 must be set equal to one or another of the eigenvalues of $\mathbb{S}$, i.e. to one or another of the zeroes of the characteristic polynomial

$$\det(\mathbb{S} - \lambda\mathbb{I}) = s_0 + s_1\lambda + s_2\lambda^2 + \cdots + s_N\lambda^N$$

and $\boldsymbol{A}$ must be proportional to the corresponding eigenvector. Familiarly, the real symmetry of $\mathbb{S}$ is by itself sufficient to insure (i) the reality of the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_N$ and (ii) the orthogonality of the eigenvectors: $\boldsymbol{A}_m \cdot \boldsymbol{A}_n = 0$ if $m \neq n$ which (after normalization) can be notated $\boldsymbol{A}_m \cdot \boldsymbol{A}_n = \delta_{mn}$. This is valuable information, but not in itself sufficient to provide *explicit descriptions* of the eigenvalues/vectors. It is, in particular, not in itself sufficient to establish that all eigenvalues are necessarily *non-negative*. Proceeding therefore (for the moment) formally, we note that the linearity of the equations of motion (5) carries with it a "principle of superposition," and are led to write

$$\begin{aligned}\boldsymbol{\varphi}(t) &= \sum_{n=1}^{N} \alpha_n \boldsymbol{A}_n e^{+i\omega_n t} + \sum_{n=1}^{N} \beta_n \boldsymbol{A}_n e^{-i\omega_n t} \\ &= \text{superposition of "normal modes"}\end{aligned}$$

with $\omega_n = \sqrt{\lambda_n}$. The complex numbers α_n and β_n are fixed by imposition of the requirements that $\boldsymbol{\varphi}(t)$ be real and that it conform to the prescribed initial data $\boldsymbol{\varphi}(0)$ and $\dot{\boldsymbol{\varphi}}(0)$; the orthonormality of the eigenvectors greatly simplifies the computational labor at this point.

The program sketched above is in fact not at all specific to crystals, but pertains generally to the "theory of small oscillations," i.e., to the classical motion of *all* particulate systems as they jiggle harmonically about points of stable equilibrium. And it admits of a great variety of alternative formulations, the relative utility of which depends upon particular features of the system in hand, and the nature of the questions uppermost in one's mind. Of these, I must be content here to sketch only one:[1]

[1] For a fairly elaborate review of the formal possiblities, see Chapters I and II of my Classical Theory of Fields (1965).

Proceeding very much in the spirit of Hamilton, let us agree to promote $\dot{\boldsymbol{\varphi}}$ to the status of an independent variable, writing $\dot{\boldsymbol{\varphi}} \equiv \boldsymbol{\chi}$. In place of (5) we can then write

$$\dot{\boldsymbol{\Phi}} = \mathbb{W}\boldsymbol{\Phi} \tag{6}$$

with

$$\boldsymbol{\Phi} = \begin{pmatrix} \boldsymbol{\varphi} \\ \boldsymbol{\chi} \end{pmatrix} \qquad \text{and} \qquad \mathbb{W} = \begin{pmatrix} \mathbb{O} & \mathbb{I} \\ -\mathbb{S} & \mathbb{O} \end{pmatrix}$$

Whereas (5) is a coupled system of N differential equations of 2^{nd} order, (6) is a $2n$-fold system of 1^{st} order, as admits therefore immediately of formal solution:

$$\boldsymbol{\Phi}(t) = e^{\mathbb{W}t}\boldsymbol{\Phi}(0)$$

The problem at this point is to assign explicit meaning to the matrix $e^{\mathbb{W}t}$. It is sometimes possible to gain useful information directly from the expansion

$$e^{\mathbb{W}t} = \sum_{n=0}^{\infty} \tfrac{1}{n!}(\mathbb{W}t)^n$$

but I propose to sketch an alternative mode of approach. By the Cauchy integral theorem

$$\tfrac{1}{2\pi i}\oint_C \tfrac{1}{z-W}e^{zt}dz = e^{Wt}$$

where C is any closed contour which envelops the singularity (simple pole) at $z = W$. We therefore expect to have

$$e^{\mathbb{W}t} = -\tfrac{1}{2\pi i}\oint_C \mathbb{R}(z)e^{zt}dz \quad \text{with} \quad \mathbb{R}(z) \equiv (\,\mathbb{W} - z\mathbb{I}\,)^{-1}$$

where C is any contour which envelops the *spectrum* of $\mathbb{W}$—the set of z-values at which $\mathbb{W} - z\mathbb{I}$ fails to be invertible. Such z-values are, of course, precisely the eigenvalues of $\mathbb{W}$. This mode of proceeding becomes useful when one is in position to produce an explicit description of $(\,\mathbb{W}-z\mathbb{I}\,)^{-1}$ in which the eigenvalues stand nakedly exposed, and can be further refined when one possesses also explicit descriptions of the associated eigenvectors.

Passage to the continuous limit by "refinement of the lattice." Suppose we had physical interest in the propagation of weak compressional waves along a wire. It is known that wires are composed of atoms, and plausible that the wire might successfully be modeled by a lattice of the design considered in the preceeding section. But it seems extravagant to invoke "atomicity" in the description of a system which to eye and instrument appears to be so continuous. We are motivated to seek a dynamical formalism which conforms more naturally to the macroscopic physics of the wire-as-we-perceive-it—a *field theory* of wires which operates in the *smooth approximation*. Such a theory (which has formal/practical interest vastly deeper, it goes without saying, than the "physics of wires") can be constructed as follows:

We embed the physical lattice within a population of similarly-structured but merely "mental" lattices. All members of the population have the same total length ℓ and the same total mass M; the number N of constituent "atoms" is, however, considered to increase without bound, with the result that both the lattice constant a and the atomic mass m drop asymptotically (i.e., in the "continuous limit") to zero. The scheme, which I call "refinement of the lattice,"

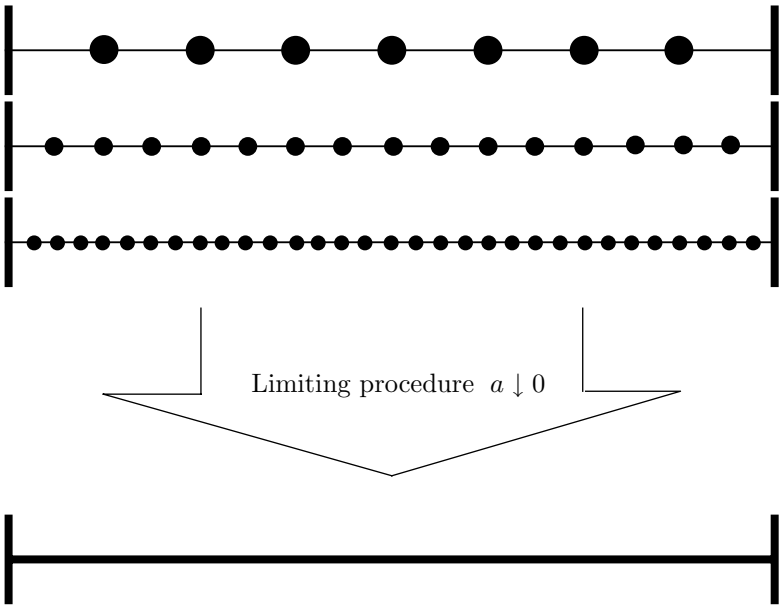


FIGURE 3: *Simple essentials of the "lattice refinement" procedure.*

is illustrated in the preceding figure. Taking the lattice constant to be our control parameter, we have

$$\left.\begin{aligned} N(a) = \frac{\ell}{a} - 1 = \frac{\ell}{a}\left(1 - \frac{a}{\ell}\right) \quad &\longrightarrow \quad \frac{\ell}{a} \\ m(a) = \frac{M}{N(a)} = \frac{M}{\ell} a \left(1 - \frac{a}{\ell}\right)^{-1} \quad &\longrightarrow \quad \mu a \end{aligned}\right\} \quad \text{for } a \ll \ell$$

where $\mu = M/\ell$ defines the *linear mass density* of the system.

As $a \downarrow 0$ the position $x_n = na$ of the n^{th} atom squeezes (for all n) ever closer to the left end of the lattice. Evidently our former practice—the ordinal enumeration of the constituent elements of our N-particle system—must be abandoned in the continuous limit. To circumvent this problem, we agree to write $\varphi(x)$ to describe the displacement (from equilibrium) of the mass element which at equilibrium resides at x $(0 \leq x \leq \ell)$. Where formerly we wrote φ_n to describe the displacement of the n^{th} atom we would now write $\varphi(x_n)$; evidently it is still possible but now no longer essential that x range on a discrete set.

In this modified notation we might, by (4), write

$$\ddot{\varphi}(x) = \frac{k}{m(a)}\big\{\varphi(x+a) - 2\varphi(x) + \varphi(x-a)\big\} \quad : \quad x = x_2, x_3, \ldots, x_{N-1}$$

to describe the motion of the typical (i.e., non-terminal) element of a discrete lattice (or "crystal"). But this equation is beautifully adapted to the formal needs of our projected "passage to the continuous limit." For we can write

$$\ddot{\varphi}(x) = \frac{k}{m(a)}a^2 \cdot \underbrace{\left\{ \frac{\frac{\varphi(x+a)-\varphi(x)}{a} - \frac{\varphi(x)-\varphi(x-a)}{a}}{a} \right\}}_{\text{becomes } \frac{\partial^2\varphi}{\partial x^2} \text{ as } a \downarrow 0}$$

and to achieve a sensible result have only to require that

$$\lim_{a\downarrow 0} \frac{ka^2}{m(a)} \sim \lim_{a\downarrow 0} \frac{ka}{\mu} = \text{a non-zero constant, call it } c^2$$

where c has necessarily the dimensionality of a "velocity" (but is not, at this point, to be associated with the "speed of light"). Evidently "regraduation of the spring constant" is a forced attribute of the lattice refinement procedure; we must have

$$k(a) = \frac{c^2\mu}{a}\left(1 + \text{inconsequential terms of order } \frac{a}{\ell}\right) \tag{7}$$

according to which the inter-atomic springs become necessarily stronger and stronger as the lattice refinement process proceeds: $k(a) \longrightarrow \infty$ as $a \downarrow 0$. This surprising development can be made intuitively intelligible by the following line of argument: Springs compose by the "law of capacitors;" for springs k_1 and k_2 connected in series one has

$$\frac{1}{k_{\text{effective}}} = \frac{1}{k_1} + \frac{1}{k_2}$$

so if one considers a spring K of length ℓ to have been assembled by connecting in series $N+1$ identical springs $k(a)$ of length $a = \ell/(N+1)$ one has

$$\frac{1}{K} = N\frac{1}{k(a)} \quad \text{with} \quad N = \frac{\ell}{a}\left(1 - \frac{a}{\ell}\right)$$

giving

$$k(a) = NK = \frac{K\ell}{a}\left(1 - \frac{a}{\ell}\right)$$

To recover (7) we have only to set $K = c^2\mu/\ell$, which we imagine to be a *constant of the refinement process*—the same in the limiting case of a compressional wire (or "string") as for the discrete crystal from which we started. Since ℓ and

μ have already been assumed to be constants of the refinement process, this amounts to a stipulation that c^2 be such a constant.

Thus is the large system (4) of coupled *ordinary* differential equations of motion seen "in the continuous limit" to go over into a *single partial* differential equation of motion:

$$\frac{1}{c^2}\frac{\partial^2\varphi}{\partial t^2} - \frac{\partial^2\varphi}{\partial x^2} = 0 \tag{8}$$

This, of course, is precisely the familiar "wave equation," of which the following are frequently-used notational variants:

$$\left\{\frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2}\right\}\varphi(x,t) = 0$$

$$\left\{\left(\frac{1}{c}\frac{\partial}{\partial t}\right)^2 - \left(\frac{\partial}{\partial x}\right)^2\right\}\varphi = 0$$

$$\left\{\tfrac{1}{c^2}\partial_t^2 - \partial_x^2\right\}\varphi = 0$$

$$\tfrac{1}{c^2}\varphi_{tt} - \varphi_{xx} = 0$$

We will encounter wave equations of many types and structures before we are done, but when one speaks of *the* wave equation one invariably has in mind either (8) or its higher-dimensional generalization

$$\left\{\tfrac{1}{c^2}\partial_t^2 - \nabla^2\right\}\varphi = 0$$

The wave equation has been demonstrated to arise in what might be called the "continuous approximation" from the theory of simple crystals, but it plays a fundamental role also in contexts—electrodynamics, for example—where no "underlying atomicity" is, so far as we are aware, present in the physics. We have learned to read in the $\partial_t^2\varphi$-term an echo of the fact that, according to Newton, *acceleration* is the kinematic variable under the direct control of F/m, and to read in the $\partial_x^2\varphi$-term an allusion to the fact harmonic *nearest-neighbor* interactions dominate the physics of crystals.[2] It is tempting to suppose that the essentials of this insight pertain to *all* natural occurances of the wave equation, and it becomes interesting in this light to recall that it is ultimately from the wave equation that we acquire an interest in the Lorentz group—an interest, that is to say, in special relativity.

Wave functions, and their relation to solutions of the lattice equations. When Newton looked to the physics of one-dimensional crystals to model the acoustic

[2] Interaction with *next* nearest neighbors would introduce ∂_x^4-terms into the associated wave equation. Generally, increased non-locality entails radically increased complication of the associated field theory. For an account of the details see C. Barnes, "The dynamics of flylines and other classical strings" (Reed College, 1992), which won for its author the APS's Apker Award.

vibrations of an air column, it was because he lacked access to a well-developed theory of partial differential equations; he found it easier to contemplate the implications of (4) than to write and study (8). We, however, are in the reverse situation. Though to do so has somewhat the nature of a digression, I look now to some of the most elementary implications of (8). I proceed in special reference to this question: To what extent do solutions of the wave equation (8) serve to clarify—and to what extent to misrepresent—the physics of real crystals ($a \neq 0$)?

Let us agree henceforth to use the term "wave function" to denote any solution of the wave equation (8), which we notate $\Box\varphi = 0$ with

$$\Box \equiv \tfrac{1}{c^2}\partial_t^2 - \partial_x^2 = \left(\tfrac{1}{c}\partial_t + \partial_x\right)\left(\tfrac{1}{c}\partial_t - \partial_x\right) \tag{9}$$

Clearly

$$\begin{aligned} \left(\tfrac{1}{c}\partial_t + \partial_x\right) f = 0 \quad &\longleftrightarrow \quad f = f(x - ct) \\ \left(\tfrac{1}{c}\partial_t - \partial_x\right) g = 0 \quad &\longleftrightarrow \quad g = g(x + ct) \end{aligned}$$

where $f(\cdot)$ and $g(\cdot)$ are any differentiable functions of a single variable. It is therefore plausible (also true!) that the most general wave function can be described

$$\begin{aligned} \varphi(x,t) &= f(x - ct) + g(x + ct) \\ &= \text{ right-running waveform} + \text{ left-running waveform} \end{aligned} \tag{10}$$

The representation (10) is, it should be noted, preserved under superposition. But it will, in general, not be directly evident to the casual eye of the person who is simply watching the motion of φ. When right and left-running waves collide they do so non-interactively—by simple superposition—and emerge from their encounter unscathed/unaltered. But in typical applications $f(\cdot)$ and $g(\cdot)$ will "sense each other's structure" (i.e., be structurally correlated) in forced consequence of imposed *boundary conditions*. Thus

$$\varphi(0,t) = 0 \quad \text{(all } t\text{)}$$

entails $g(x) = -f(-x)$, while the additional requirement

$$\varphi(\ell,t) = 0 \quad \text{(all } t\text{)}$$

would force $f(\cdot)$ to be periodic: $f(x) = f(x + 2\ell)$ for all x.

Implicit already in some preceeding remarks is the important fact that from the linearity of the wave equation it follows that wave functions are subject to a *principle of superposition*:

$$\text{wave function} + \text{wave function} = \text{wave function}$$

We stand therefore in position to consider representations of the form

$$\text{complicated wave function} = \sum \text{simple wave functions}$$

as Fourier (who in point of historical fact worked in—among others—precisely this physical context) was among the first to appreciate. Following now in Fourier's footsteps, we find it natural to write

$$f(x) = \int_{\infty}^{\infty} F(k)e^{ikx}dk \quad \text{and} \quad g(x) = \int_{\infty}^{\infty} G(k)e^{ikx}dk$$

giving

$$f(x-ct) = \int_{\infty}^{\infty} F(k)e^{ik(x-ct)}dk \tag{11.1}$$
$$= F(k)\text{-weighted superposition of right-running } \textit{harmonic waves}$$
$$g(x-ct) = \int_{\infty}^{\infty} G(k)e^{ik(x+ct)}dk \tag{11.2}$$
$$= G(k)\text{-weighted superposition of left-running } \textit{harmonic waves}$$

The harmonic waves encountered above are "simple wave functions" in the sense that they spring to our attention when we define

$$\text{phase} = kx - \omega t$$

and ask: Under what condition is $e^{i\,(\text{phase})}$ a wave function? Immediately $\omega^2 - c^2k^2 = 0$, which entails

$$\omega(k) = \pm ck$$

From

$$\tfrac{d}{dt}(\text{phase}) = k\dot{x} - \omega = 0$$

we obtain

$$\dot{x} = \text{phase velocity} = \omega/k = \pm c$$

The "rigidity" of $f(x-ct)$ is traced thus to the k-independence of the phase velocities of the "harmonic wave functions" from which, according to (11.1), it can be considered to have been assembled; the Fourier components of $f(x-ct)$ move in synchrony, and the wave is said therefore to be "non-dispersive."

Impose now the boundary conditions $\varphi(0,t) = \varphi(\ell,t) = 0$ (all t) natural to the physics of our original crystal. Spatial periodicity is then, as we have seen, enforced, and we are led at length to wave functions of the form

$$\varphi(x,t) = \sum_{n=1}^{\infty} A_n \sin k_n x \cdot e^{i\omega_n t} \tag{12}$$
$$= \text{weighted superposition of harmonic } \textit{standing waves}$$

where

$$k_n = n\pi/\ell \quad \text{and} \quad \omega_n = ck_n$$
$$= n\omega_0 \quad \text{with} \quad \omega_0 = \pi c/\ell \quad : \quad n = 1, 2, \ldots$$

We have recovered the musical physics of a clamped string—familiar to every first-year student as it was familiar in its essentials already to Pythagoras by about 500 B.C. Several points, however, deserve comment:

Pretty clearly, the "harmonic standing waves" described above are the direct continuous-limit analogs of the harmonic "normal modes" of a crystal, and they are orthonormal in this analog

$$(2/\ell)\int_0^\ell \sin k_m x \cdot \sin k_n x dx = \delta_{mn}$$

of the statement $\boldsymbol{A}_m \cdot \boldsymbol{A}_n = \delta_{mn}$. But while the vibrational frequencies natural to a string are easy to describe and infinite in number

$$\omega_n = n\omega_0 \quad \text{with} \quad n = 1, 2, \ldots$$

the frequencies natural to a crystal are difficult to describe (zeros of a high order polynomial) and finite in number. The spatial form of a standing wave is similarly easy to describe: it is sinusoidal, with

$$\text{internodal distance} = \tfrac{1}{2}\,\text{wavelength} = \tfrac{1}{n}\ell$$

while the shape of a crystaline normal mode (eigenvector of a large matrix) is relatively difficult to describe. When—reversing Newton's procedure—one looks to the physics of strings to gain insight into the physics of crystals, one gains ease of analysis, but confronts this question: To what extent does the physics of strings speak reliably—and to what extent does it misrepresent—the physics of crystals?

For the same reason that one can draw sine waves on a screen only if the wavelength significantly exceeds the pixel size, we expect a crystal to be capable of supporting only those wave forms for which

$$\text{internodal distance} \gg \text{lattice constant}$$

Since for an N-atom crystal of length ℓ the lattice constant $a = \ell/(N+1)$, we have

$$\text{internodal distance} \begin{cases} > a & \text{for} \quad n = 1, 2, \ldots, N \\ = a & \text{for} \quad n = N \\ < a & \text{for} \quad n = N+1, N+2, \ldots \end{cases}$$

and find it natural to associate only the leading N string modes—those with frequencies

$$\omega \leq \omega_{\text{cutoff}} = N\omega_0$$

—with the modes of a crystal, and to dismiss the higher-frequency modes as artifacts. How accurate is that association? Intuitively we expect it to be most reliable—both spatially and temporally—when $\omega \ll \omega_{\text{cutoff}}$, and to become increasingly deceptive as $\omega \uparrow \omega_{\text{cutoff}}$. A proper answer requires, however, that we do precisely what we have been at pains thus far to avoid—that we actually

carry to completion the program sketched on page 4. This can, in fact, be done,[3] and yields

$$\omega_n^{\text{exact}} = \omega_0 \cdot \frac{2}{\pi}(N+1)\sin\left[\frac{n}{N+1}\frac{\pi}{2}\right] \quad : \quad N = 1, 2, \ldots, N$$

Evidently $\omega_n(\text{crystal}) < \omega_n(\text{associated string})$ in all cases, and

$$\omega_n(\text{crystal}) \sim \omega_n(\text{associated string}) \quad \text{only for } n \ll N$$

as illustrated in the accompanying figure. Pretty evidently, wave motion on a

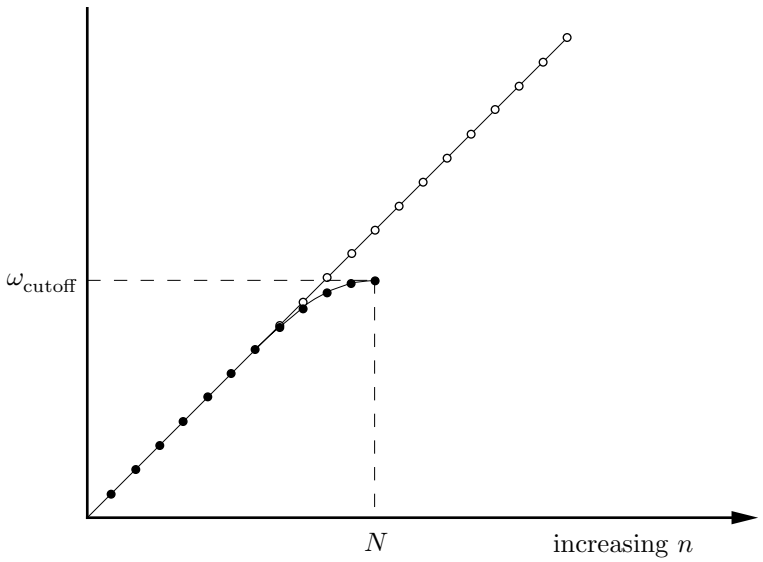


FIGURE 4: *Natural frequencies of a crystal compared with those of the associated clamped string.*

crystal is *dispersive*, and becomes (as on a string) non-dispersive only in the low-frequency limit.

If our interest attached actually to the physics of discrete N-body systems, then the moral implicit in preceeding remarks would be clear: *field-theoretic*

[3] See pp. 67–68 of U. Grenander & G. Szegö, *Toeplitz Forms and Their Applications* (1958), who exploit the fact that the $\mathbb{S}$-matrix in (5) is of such specialized structure as to comprise an instance of a "Toeplitz matrix," about which much is known. Related material can also be found in §3 of E. Montroll, "Markoff Chains, Wiener Integrals, and Quantum Theory," Comm. Pure & Appl. Math. **5**, 415 (1952). For more immediately physical discussion see W. Thompson, *Theory of Vibrations with Applications* (1972), which is the source of the result quoted on p. 50 of R. Blevins, *Formulas for Natural Frequency and Mode Shape* (1979) and reproduced here.

methods may be latently a source of striking analytic power, but must be used with cautious circumspection. In fact our interest attaches primarily to those "field-theoretic methods" themselves. For us, crystals are mere workshops, of interest primarily for such clues as they may provide concerning how we, as field-theorists, should be conducting our affairs. And so they will function in the discussion to which we now turn.

Lagrangian formulation of the wave equation. The equations of motion (4) of our one-dimensional crystal were at (3) obtained from a Lagrangian

$$L = \sum_1^N \tfrac{1}{2}m\dot{\varphi}_n^2 - \tfrac{1}{2}k\varphi_1^2 - \sum_1^{N-1} \tfrac{1}{2}k(\varphi_{n+1} - \varphi_n)^2 - \tfrac{1}{2}k\varphi_N^2 \tag{13}$$

which in notation adapted to the realities of the a-parameterized refinement process reads

$$L = \sum_1^N \tfrac{1}{2}\mu a\dot{\varphi}^2(x_n) - \tfrac{1}{2}\frac{c^2\mu}{a}\varphi^2(a) - \sum_1^{N-1} \tfrac{1}{2}\frac{c^2\mu}{a}\left[\varphi(x_n + a) - \varphi(x_n)\right]^2 - \tfrac{1}{2}\frac{c^2\mu}{a}\varphi^2(\ell - a)$$

or again (which is for our purposes more useful)

$$L = \sum_1^N \tfrac{1}{2}\mu a\dot{\varphi}^2(x_n) - \sum_1^{N-1} \tfrac{1}{2}\mu c^2 a\left[\frac{\varphi(x_n + a) - \varphi(x_n)}{a}\right]^2 - \tfrac{1}{2}\mu c^2 \cdot \tfrac{1}{a}\varphi^2(a) - \tfrac{1}{2}\mu c^2 \cdot \tfrac{1}{a}\varphi^2(\ell - a)$$

In the continuous limit $a \downarrow 0$ the dangling terms vanish in consequence of the conditions $\varphi(0) = \varphi(\ell) = 0$, and the sums become integrals; we obtain

$$L = \int_0^\ell \tfrac{1}{2}\mu c^2 \left\{ \frac{1}{c^2}\left(\frac{\partial\varphi}{\partial t}\right)^2 - \left(\frac{\partial\varphi}{\partial x}\right)^2 \right\} dx$$

It becomes natural in this light to write

$$L = \int_0^\ell \mathcal{L}dx \tag{14}$$

with

$$\mathcal{L} = \tfrac{1}{2}\mu c^2 \left\{ \tfrac{1}{c^2}\varphi_t^2 - \varphi_x^2 \right\} \tag{15}$$

and to call $\mathcal{L}$ the "Lagrangian *density*." Since μ signifies mass density, μc^2 has the physical dimensionality of an energy density, and we have

$$[\mathcal{L}] = \text{energy/length} = \text{energy density}$$

We stand now in position (i) to trace the crystaline Lagrangian (13) to its continuous limit (14), and (ii) to trace the associated system (4) of coupled

equations of motion to its continuous limit (8). It becomes natural at this point to ask: Can the field equation (8) be obtained *directly* from the Lagrangian density (15)? A little experimentation leads to the observation that

$$\tfrac{1}{c^2}\varphi_{tt} - \varphi_{xx} = 0 \quad \text{can be formulated} \quad \left\{ \frac{\partial}{\partial t}\frac{\partial}{\partial \varphi_t} + \frac{\partial}{\partial x}\frac{\partial}{\partial \varphi_x} \right\} \mathfrak{L} = 0$$

Moreover, we could in fact (since the Lagrange density of (15) displays no *explicit* φ-dependence) write

$$\left\{ \underbrace{\frac{\partial}{\partial t}\frac{\partial}{\partial \varphi_t} + \frac{\partial}{\partial x}\frac{\partial}{\partial \varphi_x}}_{\text{one such term for each independent variable}} - \frac{\partial}{\partial \varphi} \right\} \mathfrak{L} = 0 \tag{16}$$

if we wanted to maximize formal resemblance to the Lagrange equation

$$\left\{ \frac{d}{dt}\frac{\partial}{\partial \dot{\varphi}} - \frac{\partial}{\partial \varphi} \right\} L = 0$$

familiar from particle mechanics (in which context t is the *solitary* independent variable).

It is interesting to note that by nothing more complicated than a sign change ($L = T - U \longrightarrow E = T + U$) we obtain this description

$$E = \int_0^\ell \mathcal{E}\, dx \quad \text{with} \quad \mathcal{E} = \tfrac{1}{2}\mu c^2 \left\{ \tfrac{1}{c^2}\varphi_t^2 + \varphi_x^2 \right\} \tag{17}$$

of the *energy* resident on our vibrating string. That (global) energy conservation

$$\dot{E} = 0$$

is an implication of the field equation (i.e., of the equation of motion) can be established as follows:

$$\dot{E} = \mu c^2 \int_0^\ell \left\{ \tfrac{1}{c^2}\varphi_t \varphi_{tt} + \varphi_x \varphi_{xt} \right\} dx$$

But by assumption φ satisfies $\frac{1}{c^2}\varphi_{tt} - \varphi_{xx} = 0$, so

$$\begin{aligned}
&= \mu c^2 \int_0^\ell \left\{ \varphi_t \varphi_{xx} + \varphi_x \varphi_{xt} \right\} dx \\
&= \mu c^2 \int_0^\ell \frac{\partial}{\partial x}\left(\varphi_t \varphi_x \right) dx \\
&= \mu c^2 \left. \varphi_t \varphi_x \right|_0^\ell \\
&= 0 \quad \text{since } \varphi(0,t) = \varphi(\ell,t) = 0 \text{ (all } t\text{) entails } \varphi_t(0,t) = \varphi_t(\ell,t) = 0
\end{aligned}$$

Later we will be in position to discuss the deeper origins of a large population of statements of which energy conservation is but an illustrative example.

Field-theoretic formulation of Hamilton's principle. In the classical mechanics of particles it is possible to dismiss the statement

$$\delta \int L dt = 0 \quad \Longleftrightarrow \quad \left\{ \frac{d}{dt} \frac{\partial}{\partial \dot{q}_i} - \frac{\partial}{\partial q_i} \right\} L = 0 \quad : \quad i = 1, 2, \ldots, n$$

as but an elegant curiosity, for one enters into discussion of Hamilton's principle already in full command—thanks *via* Lagrange to Newton—of the equations of motion. The history of field theory supplies, however, no "Newton"—no ready-made general formulation of the equations of motion. In field theory it is, as will emerge, a generalization of Hamilton's principle which steps into the breech. The elegantly simple statement $\delta S = 0$ acquires in field theory a force and a degree of practical utility far beyond anything for which our pre-field-theoretic experience has prepared us. It becomes the central unifying principle of our subject—its workhorse.

Our objective here will be to establish the sense in which

$$\delta S = 0 \quad \Longrightarrow \quad \text{field equations}$$

but before we play chess we must put the pieces on the board. We begin by noticing that while a single field $\varphi(x,t)$ served to describe the longitudinal vibration of a clamped string, two fields—call them $\varphi_1(x,t)$ and $\varphi_2(x,t)$—would be required to describe the *transverse* vibration of such a system. And if our string had non-vanishing cross section we might find it necessary[4] to introduce yet another field $\varphi_3(x,t)$ to describe its torsional motion. Evidently a (finite) *set* of field functions

$$\varphi_1(x,t),\ \varphi_2(x,t),\ \ldots,\ \varphi_N(x,t)$$

will in the general case be required to describe the state of a distributed system, and these will, in the general case, be *dimensionally diverse*.[5]

Our string field φ was a t-dependent structure defined on a line, but in general our field systems $\varphi_1, \varphi_2, \ldots, \varphi_N$ (collectively denoted φ) will reside on manifolds of *several* dimensions. We write $x^1, x^2, \ldots, x^n$ (collectively $\boldsymbol{x}$) to refer to some specified coordinatization of the manifold. Typically we will have $n = 3$ and x^1, x^2, x^3 will refer to a *Cartesian* coordinate system, but by no means—consider the field that lives on a torus—will that be universally the case.

[4] See in this connection §6.1 "Generation of torsional waves by bow-friction forces" in L. Cremer, *The Physics of the Violin* (1981).

[5] In the preceeding example φ, φ_1 and φ_2 refer to spatial displacements, and have therefore the dimensionality of length, while the angular variable φ_3 is dimensionless.

Only rarely in particle mechanics do the physical constants which enter into the description of a system permit the formation of a "natural velocity." In field theory, on the other hand, it is rare when the available physical constants do *not* permit the formation of one or more natural velocities—constants $c_1, c_2, \ldots$ which describe (or at any rate enter into the description of) the "rapidity with which effects propagate."[6] In this respect the "velocity of light" $c = 1/\sqrt{\epsilon_0 \mu_0}$ is entirely typical, though it is atypical in that it is assigned by special relativity a preferred role which in the end has nothing special to do with light! In the presence of such a constant it becomes possible to write $x^0 = ct$, and natural in place of $\varphi(t, \boldsymbol{x})$ to adopt the still more compact notation $\varphi(x)$. This we will frequently find it convenient to do, and collaterally to write simply $\partial\varphi$ when we have in mind the entire population of first partials

$$\varphi_{\alpha,i} = \frac{\partial \varphi_\alpha}{\partial x^i} \quad \text{where} \quad \begin{cases} \alpha = 1, 2, \ldots, N \\ i \ = 0, 1, \ldots, n \end{cases}$$

At a deeper level, these opportunistic adjustments invite one to think of the fields $\varphi(x)$ not as objects that move on an n-dimensional manifold, but as objects that inhabit an $(n+1)$-dimensional spacetime.

Our chess board is now set up; it is time to play the game. As an opening move, we assume a Lagrange density of the form $\mathcal{L}(\varphi, \partial\varphi, x)$ to have been given.[7] Within the particular context provided by our clamped string system we find it natural to introduce an *action functional* by writing

$$S[\varphi(x,t)] = \int_{t_1}^{t_2} L dt \quad \text{with} \quad L = \int_0^\ell \mathcal{L}(\varphi, \partial\varphi, x) dx$$

or, more compactly,

$$S[\varphi(x,t)] = \iint_{\mathcal{R}} \mathcal{L}(\varphi, \partial\varphi, x)\, dt dx$$

where $\mathcal{R}$ refers to the "rectangular box" in spacetime defined $t_1 \leq t \leq t_2$, $0 \leq x \leq \ell$. By straightforward extension, we take $\mathcal{R}$ to be an *arbitrary* domain (or "bubble") in $n+1$-dimensional spacetime and agree to let

$$S_{\mathcal{R}}[\varphi] = \int_{\mathcal{R}} \mathcal{L}(\varphi, \partial\varphi, x) dt dx^1 \cdots dx^n \tag{18}$$

serve in the general case to define the "action functional relative to $\mathcal{R}$" of the field system $\mathcal{L}(\varphi, \partial\varphi, x)$: $\varphi = \{\varphi_1, \varphi_2, \ldots, \varphi_N\}$. In the following figure I have

[6] The non-relativistic quantum mechanics of a particle—thought of as a classical field theory—is in this respect the great exception.

[7] Note the assumed absence of arguments of the type $\partial\partial\varphi, \partial\partial\partial\varphi, \ldots$ Such terms, were we proceeding from a crystaline model, would reflect the presence of next-nearest-neighbor and even more remote interactions. Evidently we proceed subject to a tacit *locality* assumption.

attempted to represent the geometrical image one has in mind when one draws upon the fundamental definition (18).

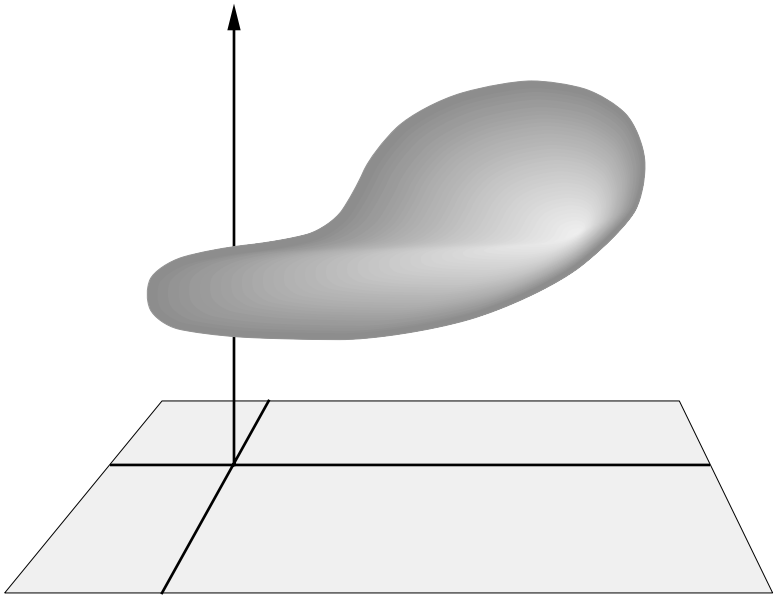


FIGURE 5: *A "bubble" $\mathcal{R}$ in the $(n+1)$-dimensional spacetime inhabited by the field system $\{\varphi_1, \varphi_2, \ldots, \varphi_N\}$*

We look to the leading-order response $S_{\mathcal{R}} \longrightarrow S_{\mathcal{R}} + \delta S_{\mathcal{R}}$ of the action functional $S_{\mathcal{R}}[\varphi]$ to hypothetical variation $\varphi \longrightarrow \varphi + \delta\varphi$ of the field system, subject to the explicit stipulation that

$$\delta\varphi = 0 \quad \text{on the boundary } \partial\mathcal{R} \text{ of } \mathcal{R}$$

Writing[8]

$$\begin{aligned}\delta S_{\mathcal{R}}[\varphi] &= S_{\mathcal{R}}[\varphi + \delta\varphi] - S_{\mathcal{R}}[\varphi] \\ &= \int_{\mathcal{R}} \big\{\mathcal{L}(\varphi + \delta\varphi, \partial\varphi + \delta\partial\varphi, x) - \mathcal{L}(\varphi, \partial\varphi, x)\big\} dx\end{aligned}$$

[8] I adopt here and henceforth the abbreviation

$$dx = dt dx^1 \cdots dx^n = \tfrac{1}{c} dx^0 dx^1 \cdots dx^n$$

where c has been selected from the population of velocities $c_1, c_2, \ldots$ natural to the system in hand. The latter variant, though almost always available in principle, will acquire special naturalness and utility in connection with the theory of *relativistic* classical fields.

we (by Taylor expansion of the integrand) obtain

$$\delta S_{\mathcal{R}}[\varphi] = \int_{\mathcal{R}} \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} \delta\varphi_\alpha + \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} \delta\varphi_{\alpha,i} \right\} dx$$

where $\sum_\alpha$ and $\sum_i$ are understood. Pretty clearly

$$\delta\varphi_{\alpha,i} = \partial_i \left(\delta\varphi_\alpha\right)$$

so, integrating by parts, we have

$$\delta S_{\mathcal{R}}[\varphi] = \int_{\mathcal{R}} \delta\varphi_\alpha \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} - \frac{\partial}{\partial x^i} \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} \right\} dx + \int_{\mathcal{R}} \frac{\partial}{\partial x^i} \left(\delta\varphi_\alpha \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} \right) dx$$

But the second of the integrals on the right has the structure $\int_{\mathcal{R}} (\partial_i \mathcal{A}^i) dx$, and by the divergence theorem[9] $\int_{\mathcal{R}} (\partial_i \mathcal{A}^i) dx = \int_{\partial\mathcal{R}} \mathcal{A}^i d\sigma_i$ so

$$\int_{\mathcal{R}} \frac{\partial}{\partial x^i} \left(\delta\varphi_\alpha \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} \right) dx = \int_{\partial\mathcal{R}} \left(\delta\varphi_\alpha \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} \right) d\sigma_i$$

which vanishes since, by assumption, $\delta\varphi = 0$ on the boundary $\delta\mathcal{R}$ of $\mathcal{R}$. Therefore

$$\delta S_{\mathcal{R}}[\varphi] = \int_{\mathcal{R}} \delta\varphi_\alpha \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} - \frac{\partial}{\partial x^i} \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} \right\} dx \tag{19}$$

By *Hamilton's Principle*, φ will be "dynamical" if and only if

$$\delta S_{\mathcal{R}}[\varphi] = 0 \quad \text{for all regions } \mathcal{R} \text{ and all variations } \delta\varphi \tag{20}$$

It follows from (19) that if the field system φ conforms to Hamilton's Principle then the field functions $\varphi_1, \varphi_2, \ldots, \varphi_N$ have necessarily to be solutions of the *field equations*

$$\left\{ \frac{\partial}{\partial \varphi_\alpha} - \frac{\partial}{\partial x^i} \frac{\partial}{\partial \varphi_{\alpha,i}} \right\} \mathcal{L}(\varphi, \partial\varphi, x) = 0 \qquad \alpha = 1, 2, \ldots, N \tag{21}$$

These comprise an N-fold system of coupled second-order partial differential equations. Equations (21) are of a form which was anticipated already at (16), and can in a more explicit notation be expressed

$$\frac{\partial \mathcal{L}}{\partial \varphi_\alpha} - \frac{\partial^2 \mathcal{L}}{\partial \varphi_\beta \partial \varphi_{\alpha,i}} \varphi_{\beta,i} - \frac{\partial^2 \mathcal{L}}{\partial \varphi_{\beta,j} \partial \varphi_{\alpha,i}} \varphi_{\beta,ij} - \boldsymbol{\partial}_i \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} = 0 \tag{22}$$

where $\sum_\alpha$, $\sum_i$ and $\sum_j$ are understood, and where $\boldsymbol{\partial}_i$ looks only to the *explicit* x-dependence of $\mathcal{L}(\cdot, \cdot, x)$. The field equations will, in the general case, be non-linear.

[9] Which is to say, by Gauß' theorem—by a particular instance of Stokes' theorem. Here $d\sigma_i$ signifies an outer-directed surface element.

Gauge freedom in the construction of the Lagrangian. The Lagrangian density $\mathcal{L}$ plays in field theory precisely the "system characterizer" role which in particle mechanics is played by the Lagrangian, $L(q, \dot{q}, t)$. The

$$\text{field system} \quad \longleftrightarrow \quad \text{Lagrangian density}$$

association is, however—like its particle mechanical counterpart—non-unique. If $\mathcal{L}$ and $\mathcal{L}'$ stand in the relation

$$\mathcal{L}'(\varphi, \partial\varphi, x) = \mathcal{L}(\varphi, \partial\varphi, x) + \partial_k \mathcal{G}^k(\varphi, x)$$

then they *give rise to identical field equations*, for the simple reason that (as can be shown by explicit calculation)

$$\left\{ \frac{\partial}{\partial \varphi_\alpha} - \frac{\partial}{\partial x^i} \frac{\partial}{\partial \varphi_{\alpha,i}} \right\} \partial_k \mathcal{G}^k(\varphi, x) = 0 \quad \text{identically, for all } \mathcal{G}^k(\varphi, x)$$

Insight into the origin of this important fact follows from the observation that the gauge transformation

$$\mathcal{L} \longrightarrow \mathcal{L}' = \mathcal{L} + \partial_k \mathcal{G}^k \tag{23}$$

induces

$$\begin{aligned} S_{\mathcal{R}} = \int_{\mathcal{R}} \mathcal{L} dx \quad \longrightarrow \quad S_{\mathcal{R}}{}' &= \int_{\mathcal{R}} \mathcal{L}' dx \\ &= S_{\mathcal{R}} + \int_{\mathcal{R}} \partial_k \mathcal{G}^k dx \\ &= S_{\mathcal{R}} + \int_{\partial\mathcal{R}} \mathcal{G}^k d\sigma_k \quad \text{by the divergence theorem} \end{aligned}$$

and that the final term—the boundary term—is, for the purposes of Hamilton's Principle, invisible.

Non-uniqueness—gauge freedom—entails that the Lagrangian density $\mathcal{L}$ is, its formal importance notwithstanding, not itself directly "physical." It is, in this respect, reminiscent of the potentials $U(x)$ of particle mechanics, which are determined only up to gauge transformations of the form

$$U(x) \longrightarrow U'(x) = U(x) + \text{constant}$$

In the laboratory one measures not potentials, but only such gauge-invariant constructs as (for example) potential *differences*. Gauge freedom imposes upon us an obligation frequently to attach (at least tacitly) tedious qualifications to statements we might prefer to keep sharply simple. For example: if $\mathcal{L}$ depends *quadratically* upon φ and $\partial\varphi$ then (see again (22)) the associated field equations will be *linear*. It would, however, be incorrect to state that "linearity implies

the quadraticity" of $\mathcal{L}$, for to $\mathcal{L}$ we could aways add a quadraticity-breaking gauge term.

We have now in hand, in the field equations (21), the general field-theoretic proposition of which (16) provided our first hint. We turn now to a discussion the objective of which will be to establish in similar generality the origins of a population of important propostions of which (17) is the precursor. We turn, in short, to a discussion of "Hamilton's Principle, Part II"—i.e., of Noether's Theorem in its original (classical field-theoretic) setting.

Field-theoretic formulation of Noether's Theorem. Emmy Noether (1882–1935) is today remembered by mathematicians primarily for the importance of her contributions to algebraic number theory, especially to the theory of ideals and to several aspects of the theory of invariants. But her name will live forever among physicists for the work which is our present subject matter. I digress to sketch the soil from which that work sprang.

Noether did the first semester of what we would today call graduate study at the University of Göttingen (winter term 1903–1904), where she audited[10] lectures by (among others) Hermann Minkowski, Felix Klein and David Hilbert. She then returned to Erlangen, where her father was a professor, where Felix Klein (1849–1925) had in his inaugural lecture (1872) propounded the influential "Erlangen Program"[11] which held the *group* and *invariance* concepts to be among the central organizing principles of mathematical (and also physical) thought, and where Paul Gordon (1837–1912) became her mentor. Research publications by E. Noether began appearing in 1907. Meanwhile...

Einstein was at the sublime height of his powers during the "miracle decade" 1905–1915, and his work—especially that relating to the development of general relativty—attracted the close attention of mathematicians, especially (at Göttingen) of Klein and Hilbert. In 1915 Noether—though unable because of her sex to obtain either an advanced degree or a paid teaching position—returned to Göttingen, where she soon became a kind of unofficial assistant to Klein and especially Hilbert. In November of 1915 Noether wrote to a friend back in Erlangen that "the theory of invariants is the thing here now; even the physicist Hertz[12] is studying [the subject]; Hilbert plans to lecture next week about his ideas on Einstein's differential invariants, and so our crowd had better be ready." To another friend she reported that she and collaborators were carrying out calculations of the most difficult kind for Einstein "although none of us understands what they are for." Klein remarks in a letter to Hilbert

[10] Women were still at the time permitted to audit, but not to enroll as students in advanced courses of study.

[11] For a good discussion of the historical impact of the Erlangen Program see E. Bell, *Development of Mathematics* (1945), pp. 442–453.

[12] The reference is to Gustav Hertz (1887–1975), later to acquire fame for his participation in the celebrated "Franck–Hertz experiment." The theoretically astute Heinrich Hertz certainly would have had interest in such material, but had died already in 1894.

that "Noether is continually advising me in my projects, and it is really through her that I have become competent in the subject..." Hilbert makes reference in his response to "Emmy Noether, whom I called upon to help me with such questions as my theorem on the conservation of energy..." It was Noether's effort to be "helpful" in precisely that connection which led to the development of "Noether's Theorem."

By 1919 it had finally become possible (owing to a change in German law; manpower had become short in a Germany at war) for a woman to earn an advanced degree, and to hold a university appointment. On 4 June 1919—six days after A. S. Eddington had obtained solar eclipse data in agreement with Einstein's prediction of the bending of star light—Noether stood before a mathematical faculty which included R. Courant, P. Debye, Hilbert, Klein, E. Landau, L. Prandtl and W. Voigt to deliver her *Habilitation* lecture. She had been active in many areas during her years at Göttingen, and could have spoken on a wide variety of topics. But she chose in fact to speak on research which she had published already in 1918 under the title "Invariante Variationsprobleme." Concerning that work she wrote at the time as follows:

> "The last...of the works to be mentioned here concern differential invariants and variational problems and in part are an outgrowth of my assistance to Klein and Hilbert in their work on Einstein's general theory of relativity...The [paper], which I designated as my *Habilitation* thesis, deals with arbitrary finite or infinite continuous groups, in the sense of Lie, and discloses what consequences it has for a variational problem to be invariant with respect to such a group. The general results contain, as special cases, the theorems on first integrals as they are known in mechanics; furthermore, the conservation theorems and the interdependences among the field equations in the theory of relativity—while, on the other hand, the converses of these theorems are also given..."

Short biography does invariable violence to the always-intricate facts of the matter. For those I must refer you, dear reader, to the relevant literature,[13] which anticipate, I think likely to make a lasting impression upon you. Here my objective has been simply to suggest that the work for which Noether's name will forever be remembered by physicists is work which is in fact clearly consonant with the principal themes evident in the larger body of her mathematical work. It very cleverly exploits and enshrines a little constellation of ideas which were very much in the air—at Göttingen and elsewhere—during

[13] My principal source has been A. Dick's *Emmy Noether* (1981), but see also *Emmy Noether: A Tribute to Her Life and Work* (edited by J. Brewer & M. Smith, and published in that same centennial year) and the deeply informed and sensitively written obituary by Hermann Weyl which can be found at p. 425 in Volume III of his *Gesammelte Adhandlugen* (1968). The circumstances associated specifically with the development of Noether's Theorem are discussed on p. 431.

the first years of this century, and which have assumed ever greater importance as the century has matured.

Noether's Theorem emerges fairly spontaneously when two ideas are (so to speak) "rubbed against each other." The first of those has to do with the concept of "dynamical action," the other with the concept of "parameterized map." We consider them in that order:

Let $\mathfrak{L}(\varphi, \partial\varphi, x)$ be given, and let $\varphi(x)$ be some *solution* of the associated field equations. We agree to write $\varphi_{\text{dynamical}}(x)$ when we wish to emphasize that it is *such* field functions—"dynamical" field functions—that we have in mind. The phrase "dynamical action" refers then to constructions of the form

$$S_{\mathcal{R}}[\varphi_{\text{dynamical}}(x)]$$

Easy enough... yet complex enough to conceal some deep mysteries, as comes quickly to light when one looks to the corresponding construct in ordinary particle mechanics. Consider $L(q, \dot{q}, t)$ to be given, and take $q_{\text{dynamical}}(t)$ to be a solution of the associated equations of motion which conforms to endpoint conditions

$$q_{\text{dynamical}}(t) = \begin{cases} q_1 & \text{when} \quad t = t_1 \\ q_2 & \text{when} \quad t = t_2 \end{cases}$$

Familiarly, the action functional $S[q(t)]$, when evaluated at $q(t) = q_{\text{dynamical}}(t)$, becomes a *function of the endpoint data*:

$$S[q_{\text{dynamical}}(t)] = S(q_2, t_2; q_1, t_1)$$

But while initial data

$$q(t_1) = q_1 \quad \text{and} \quad \dot{q}(t_1) = v_1$$

generally is sufficient to determine $q(t) = q_{\text{dynamical}}(t)$ uniquely, endpoint data generally is not; evidently the dynamical action function must, in the general case, be multi-valued. This important fact we attribute to the circumstance that while statements of the form

$$S[q(t)] = \text{extremum}$$

are global in character (and might for that reason be expected to admit in most cases of a unique solution), Hamilton's Principle $\delta S[q(t)] = 0$ imposes only a *local* condition on the trajectory $q(t)$. Returning in this light to classical field theory, we expect to have

$$S_{\mathcal{R}}[\varphi_{\text{dynamical}}(x)] = \text{some function } S(\varphi(\partial\mathcal{R})) \text{ of prescribed } \textit{boundary data}$$

Moreover, we expect $S(\varphi(\partial\mathcal{R}))$ to be in the general case *multi-valued*. But it is by no means obvious that there even *exists* a $\varphi_{\text{dynamical}}(x)$ which conforms to arbitrarily prescribed boundary data $\varphi(\partial\mathcal{R})$, and to resolve such an issue one would have to enter distractingly far into the general theory of partial

differential equations. Happily, we can proceed formally in total ignorance of such theory, and address such matters on a case by case basis as specific occasions arise.

Turning now to the concept of "parameterized map" as it enters into Noether's train of thought... the simplest manifestation of basic idea emerges naturally as soon as one agrees to look upon rotations, translations, dilations, curvilinear deformations and other such point-to-point transformations as "flows" achieved by specification of one or more continuously variable "control parameters" $\omega = \{\omega^1, \omega^2, \ldots, \omega^\nu\}$. More concretely, let a coordinate system be inscribed on the $(n+1)$-space inhabited by our field system, let x and X signify the coordinates of a point and its image, write

$$\mathcal{T}_\omega \quad : \quad x \longrightarrow X(x;\omega) \tag{24}$$

and (though it entails the exclusion of such otherwise unexceptionable—and frequently important—transformations as reflections and projections) agree to look henceforth only to cases in which the set $\mathcal{T}=\{\mathcal{T}_\omega\}$ has these properties:

- *compositional closure*: For every pair $\{\omega_1, \omega_2\}$ there exists an $\omega(\omega_1, \omega_2)$ such that $\mathcal{T}_{\omega_2}\mathcal{T}_{\omega_1} = \mathcal{T}_{\omega(\omega_1,\omega_2)}$
- *existence of an identity*: There exists within $\mathcal{T}$ an element $\mathcal{T}_{\omega_0}$ such that $X(x;\omega_0) = x$ for all x ; we henceforth assume the parameterization to have been rigged in such a way as to achieve $\omega_0 = 0$, and write $\mathcal{T}_0$ to denote the identity transformation.
- *existence of an inverse*: For every ω there exists an $\Omega(\omega)$ such that $\mathcal{T}_\Omega\mathcal{T}_\omega = \mathcal{T}_0$; in other words, $X(X(x;\omega);\Omega(\omega)) = x$ for all x and all ω.
- *associativity*: $\mathcal{T}_{\omega_3}(\mathcal{T}_{\omega_2}\mathcal{T}_{\omega_1}) = (\mathcal{T}_{\omega_3}\mathcal{T}_{\omega_2})\mathcal{T}_{\omega_1}$

We are brought thus to the notion of a "continuous group of transformations," of which $\omega(\omega_1, \omega_2)$ is, in effect, the "group multiplication table." The theory of such groups—"Lie groups," as they are called—was (together with the details of a great many illustrative applications) worked out almost single-handedly by the Norwegian mathematician Sophus Lie (1842-1899) during the 1870's and 1880's.[14] Fundamental to the theory of Lie groups is the insight that finite transformations can be built up by iteration of infinitesimal ones; the structure of a Lie group is *latent already in its structure in the infinitesimal neighborhood of the identity.* Returning in this light to (24)—which, as we have rigged things, reduce to description of the map as experienced in the

[14] Lie and the precocious Klein (seven years his junior) had been students together at Göttingen. Klein was initially Lie's collaborator, and stood always ready to lend him support and encouragement. And Klein was, as I have remarked, one of Noether's primary mentors; the ideas here at work were therefore entirely natural to her. For a good brief account of Lie's work in its original setting (unified theory of differential equations), see Chapter V of E. Ince's classic *Ordinary Differential Equations* (1926).

neighborhood of the point $\omega = 0$ in parameter space—we have

$$\mathcal{T}_{\delta\omega} \quad : \quad x \longrightarrow X(x;\delta\omega) = x + \delta_\omega x$$

with

$$\delta_\omega x = \sum_{r=1}^{\nu} \mathcal{X}_r(x)\delta\omega^r$$

where the functions $\mathcal{X}_r(x)$—called "structure functions" because it is they which account ultimately for the distinctive structure of the particular Lie group in hand—can, in the notation natural to the finite transformation (24), be described

$$\mathcal{X}_r(x) = \left.\frac{\partial X(x;\omega)}{\partial \omega^r}\right|_{w=0}$$

The (infinitesimal) "parameterized maps" (my terminology) contemplated by Noether appropriate, but at the same time enlarge upon, the root idea sketched above. The map $\mathcal{T}_{\delta\omega}$ is understood by Noether to be "bipartite," in this sense: it sends spacetime points to new spacetime points (in precisely the manner described above), and—simultaneously but quite independently—it *adjusts the functional structure* of the field functions $\varphi(x)$:

$$\mathcal{T}_{\delta\omega} \quad : \quad \begin{cases} x \longrightarrow X(x\,;\delta\omega) = x + \delta_\omega x \\ \varphi(x) \longrightarrow \Phi(X;\delta\omega) = \varphi(x) + \delta_\omega\varphi(x) \end{cases} \tag{25}$$

where (installing all indices, but surpressing a $\sum_r$)

$$\delta_\omega x^i = \mathcal{X}_r^i(x)\delta\omega^r \tag{26.1}$$

$$\delta_\omega\varphi_\alpha(x) = \Phi_{\alpha r}(x)\delta\omega^r \tag{26.2}$$

The field variation $\delta_\omega\varphi_\alpha(x)$ derives, as emphasized above, from two distinct sources, and those contributions are (since we are working in lowest order) additive; we have

$$\begin{aligned} \delta_\omega\varphi_\alpha(x) = {} & \text{contribution from variation of } \textit{argument} \\ & + \text{contribution from variation of } \textit{functional form} \end{aligned}$$

Since the former can be described $\varphi_{\alpha,i}\delta_\omega x^i$ we can notate the preceeding disentanglement as follows:

$$\begin{aligned} \delta_\omega\varphi_\alpha &= \varphi_{\alpha,i}\delta_\omega x^i + \left\{\Phi_{\alpha r}\delta\omega^r - \varphi_{\alpha,i}\delta_\omega x^i\right\} \\ &= \varphi_{\alpha,i}\delta_\omega x^i + \underbrace{\left\{\Phi_{\alpha r} - \varphi_{\alpha,i}\mathcal{X}_r^i\right\}}\delta\omega^r \\ &\qquad\qquad = \Delta_\omega\varphi_\alpha \end{aligned} \tag{27.1}$$

Similarly

$$\delta_\omega\varphi_{\alpha,i} = \varphi_{\alpha,ij}\delta_\omega x^j + \Delta_\omega\varphi_{\alpha,i} \quad \text{with} \quad \Delta_\omega\varphi_{\alpha,i} = \left(\Delta_\omega\varphi_\alpha\right)_{,i} \tag{27.2}$$

which serves to disentangle the variations of the various field derivatives.

Armed as we are with some understanding of the concepts of "dynamical action" and "parameterized map," we are in position now at last to put those ideas into the same pot and stir; we will find that Noether's Theorem emerges (as I have claimed) "fairly spontaneously," but not without the exercise of some trickery. We look to the description of

$$\delta_\omega S_{\mathcal{R}}[\varphi] = S_{\mathcal{R}+\delta\mathcal{R}}[\Phi] - S_{\mathcal{R}}[\varphi]$$

subject to the assumption that φ is a solution of the field equations. Writing

$$S_{\mathcal{R}+\delta\mathcal{R}}[\Phi] = \int_{\mathcal{R}+\delta\mathcal{R}} \mathfrak{L}(\Phi(X), \ldots) dX$$

we have

$$= \int_{\mathcal{R}} \mathfrak{L}(\Phi(X(x)), \ldots) \left| \frac{\partial X}{\partial x} \right| dx$$

after the indicated change of variables.[15] Therefore (introducing a term at the beginning only to subtract it again at the end)

$$\delta_\omega S_{\mathcal{R}}[\varphi] = \int_{\mathcal{R}} \left\{ \mathfrak{L}(\Phi(X(x)), \ldots) - \mathfrak{L}(\varphi(x), \ldots) \right\} dx + \int_{\mathcal{R}} \mathfrak{L}(\Phi(X(x)), \ldots) \left\{ \left| \frac{\partial X}{\partial x} \right| - 1 \right\} dx$$

Expansion of the Jacobian (use $\det(\mathbb{I} + \epsilon\mathbb{M}) = 1 + \epsilon \mathrm{tr}\mathbb{M} + \cdots$) gives

$$\left\{ \left| \frac{\partial X}{\partial x} \right| - 1 \right\} = \frac{\partial}{\partial x^k} (\delta_\omega x^k) + \cdots$$

Since this expression is itself of first order, and we are working only in first order, we can in the second integral replace $\mathfrak{L}(\Phi(X(x)), \ldots)$ by its zeroth order approximation $\mathfrak{L}(\varphi(x), \ldots)$, giving

$$\delta_\omega S_{\mathcal{R}}[\varphi] = \int_{\mathcal{R}} \left\{ \mathfrak{L}(\Phi(X(x)), \ldots) - \mathfrak{L}(\varphi(x), \ldots) \right\} dx + \int_{\mathcal{R}} \mathfrak{L}(\varphi(x), \ldots) \frac{\partial}{\partial x^k} (\delta_\omega x^k) dx$$

Turning our attention now to the first of the integrals in the preceeding equation, we use

$$\Phi_\alpha(X) = \varphi_\alpha(x) + \delta_\omega \varphi_\alpha(x) \quad \text{and} \quad \Phi_{\alpha,k}(X) = \varphi_{\alpha,k}(x) + \delta_\omega \varphi_{\alpha,k}(x)$$

to obtain

[15] Noting that every X in $\mathcal{R} + \delta\mathcal{R}$ is the image under $\mathcal{T}_{\delta\omega}$ of an x in $\mathcal{R}$, we have elected to adopt the latter as our variables of integration. This has the effect of making both integrals range on the same domain.

$$\int_{\mathcal{R}} \{\mathcal{L}(\Phi(X(x)),\ldots) - \mathcal{L}(\varphi(x),\ldots)\}dx$$
$$= \int_{\mathcal{R}} \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha}\delta_\omega \varphi_\alpha + \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,j}}\delta_\omega \varphi_{\alpha,j} + \frac{\boldsymbol{\partial}\mathcal{L}}{\boldsymbol{\partial} x^k}\delta_\omega x^k \right\} dx$$

which by (27) becomes

$$= \int_{\mathcal{R}} \left\{ \frac{\partial \mathcal{L}}{\partial \varphi_\alpha}\Delta_\omega \varphi_\alpha + \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,j}}\frac{\partial}{\partial x^j}(\Delta_\omega \varphi_\alpha) \right.$$
$$\left. + \underbrace{\left[\frac{\partial \mathcal{L}}{\partial \varphi_\alpha}\varphi_{\alpha,k} + \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,j}}\varphi_{\alpha,jk} + \frac{\boldsymbol{\partial}\mathcal{L}}{\boldsymbol{\partial} x^k} \right]}_{= \frac{\partial \mathcal{L}}{\partial x^k}} \delta_\omega x^k \right\} dx$$

so after some slight manipulation we obtain

$$= \int_{\mathcal{R}} \left\{ \delta_\omega x^k \frac{\partial \mathcal{L}}{\partial x^k} + \underbrace{\left[\frac{\partial \mathcal{L}}{\partial \varphi_\alpha} - \frac{\partial}{\partial x^j}\frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,j}} \right]}_{0} \Delta_\omega \varphi_\alpha + \frac{\partial}{\partial x^k}\left(\frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}}\Delta_\omega \varphi_\alpha \right) \right\} dx$$

Here the expression internal to the square bracket *vanishes by virtue of our assumption that φ is dynamical*. Combining this result with that achieved near the bottom of the preceeding page, we obtain

$$\delta_\omega S_{\mathcal{R}}[\varphi] = \int_{\mathcal{R}} \frac{\partial}{\partial x^k}\left[\frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}}\Delta_\omega \varphi_\alpha + \mathcal{L}\delta_\omega x^k \right] dx$$

Drawing finally upon (26.1) and (27.1), we obtain Noether's Theorem:

$$\delta_\omega S_{\mathcal{R}}[\varphi_{\text{dynamical}}] = \sum_{r=1}^{\nu} \delta\omega^r \cdot \int_{\mathcal{R}} \left(\partial_k J_r^k\right) dx = \sum_{r=1}^{\nu} \delta\omega^r \cdot \int_{\partial\mathcal{R}} J_r^k d\sigma_k \tag{28}$$

with

$$J_r^k = \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}}\left\{ \Phi_{\alpha r} - \varphi_{\alpha,i}\mathcal{X}_r^i \right\} + \mathcal{L}\mathcal{X}_r^k \tag{29}$$

where $\sum_\alpha$ and $\sum_k$ are understood.

Equation (29) can—quite naturally, in view of the construction $\partial_k J_r^k$ which made unbidden claim to our attention at (28)— be considered to describe the k-indexed *components* of an object $\boldsymbol{J}_r$. One such object—one such "Notherian current"—is associated with each of the parameters ω^r which enter into the description of the map $\mathcal{T}$. One can write out the the explicit description

$$J_r^k = J_r^k(\varphi, \partial\varphi, x)$$

of such a $\boldsymbol{J}_r$ as soon as one is in possession of (i) the Lagrangian density $L(\varphi, \partial\varphi, x)$ characteristic of the system in hand, and (ii) the structure functions $\mathcal{X}^i_r(x)$ and $\Phi_{\alpha r}(\varphi, x)$ characteristic of the map. The question, however, remains: what is such knowledge good for?

General considerations relating to the application of Noether's Theorem. We came at (28) to a conclusion of which

$$\delta_\omega S_{\mathcal{R}}[\varphi] = \sum_r \delta\omega^r \cdot \int_{\mathcal{R}} \text{div}\boldsymbol{J}_r dx = \sum_r \delta\omega^r \cdot \int_{\partial\mathcal{R}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} \tag{30}$$

provides a picturesque abbreviation. This is a result of charming simplicity, but it is for the power of its immediate implications that it is celebrated. Suppose, for example, that it could on some grounds be asserted that

$$\delta_\omega S_{\mathcal{R}}[\varphi] = 0 \quad \text{for all bubbles } \mathcal{R} \text{ and all variations } \delta\omega \tag{31}$$

It would then follow that

$$\partial_k J^k_r = 0 \quad : \quad r = 1, 2, \ldots, \nu \tag{32}$$

These are "continuity equations," statements of the form

$$\tfrac{\partial}{\partial t}(\text{density}) + \boldsymbol{\nabla}\cdot(\textbf{flux}) = 0$$

What we have in (32) is an ν-fold set of *conservation laws*.

Insofar as (31) $\Longrightarrow$ (32), Noether's Theorem serves to provide a particularly precise and powerful formulation of the connection between *symmetries* (of the dynamical action) on the one hand, and *conservation laws* on the other. It derives its power in part from the fact that it formulates the association

$$\text{symmetry} \quad \longleftrightarrow \quad \text{conservation}$$

in terms which are rooted in a variational principle, and which are, therefore, essentially coordinate-free. When a new conservation law has been discovered (experimentally, let us say), it becomes urgent in this light to undertake a search for the underlying symmetry, and when such a symmetry is discovered it is difficult to resist the conclusion that one has discovered something "deep."[16]

It is useful to notice that Noether's Theorem gives rise to currents $\boldsymbol{J}$ which tend generally to be "interesting" to precisely the degree that the associated map is interesting—whether or not $\boldsymbol{J}$ happens to be in fact conserved. Energy, momentum, angular momentum... are, interesting (because frequently useful) physical constructs even in contexts where they are not conserved. Special

[16] Such searching may, however, prove futile. Contrary to a widely-held belief, there exist conservation laws which do *not* have their origin in invariance properties of the dynamical action.

interest attaches (but not exclusively) to maps which embody *isometries of the spacetime manifold*.

One should bear in mind that conservation laws—whatever the "symmetry considerations" that may have been that led to their discovery—have ultimately this status: they are *implications of the equations of motion*. This is true even when (as in relativistic field theory) the symmetry is one which has been intentionally "built into" the field equations. By way of illustration, consider the simple "translational map"

$$\mathcal{T}^{\text{translation}}_{\delta\omega} \quad : \quad \begin{cases} x^i \longrightarrow X^i(x\,;\delta\omega) = & x^i + \delta\omega^i \\ \varphi_\alpha(x) \longrightarrow \Phi_\alpha(X;\delta\omega) = & \varphi_\alpha(x) \end{cases} \tag{33}$$

Comparison with (26) shows the associated structure functions to be given by

$$\mathcal{X}^i_j = \delta^i_j \quad \text{and} \quad \Phi_{\alpha j} = 0$$

We are led thus from (29) to expressions of the design

$$J^k_j = \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}} \Big\{ 0 - \varphi_{\alpha,i}\delta^i_j \Big\} + \mathcal{L}\delta^k_j$$

which, in respect for entrenched tradition, we agree to notate

$$S^k{}_j = \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}} \varphi_{\alpha,j} - \mathcal{L}\delta^k{}_j \tag{34}$$

and to call the "stress-energy tensor."[17] Does (33) describe in fact asymmetry, in the sense $\delta_\omega S_{\mathcal{R}}[\varphi] = 0$, of the dynamical action? Is it in fact the case that $\partial_k S^k{}_j = 0$? By calculation

$$\partial_k S^k{}_j = \left[\frac{\partial}{\partial x^k} \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}} \right] \varphi_{\alpha,j} + \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}} \varphi_{\alpha,kj} - \frac{\partial \mathcal{L}}{\partial \varphi_\alpha} \varphi_{\alpha,j} - \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,i}} \varphi_{\alpha,ij} - \frac{\boldsymbol{\partial} \mathcal{L}}{\boldsymbol{\partial} x^k}$$

Since the second term cancels the fourth (trivially), and the first cancels the third *in consequence of the equations of motion*, we have

$$\begin{aligned} \partial_k S^k{}_j &= -\frac{\boldsymbol{\partial} \mathcal{L}}{\boldsymbol{\partial} x^k} \\ &= 0 \quad \text{if and only if } \mathcal{L} \text{ has no } \textit{explicit } x\text{-dependence} \end{aligned}$$

[17] It is, of course, entirely natural to assign particularized names/notations to Noetherian currents which—as here—derive from particularized assumptions. But when one attaches the word "tensor" to an object one is not simply alluding to its indicial decorations; one is making a statement concerning the explicit *transformation properties* of the object in question. We ourselves have yet to discuss the transformation properties of $S^k{}_j$.

In all applications of Noether's Theorem one stands with one foot planted in particularities of the map, and the other in particularities of the system—that is, of the Lagrange density $\mathfrak{L}$ which serves to describe the system. In the preceeding discussion $\mathfrak{L}$ remained unspecified at (34), and we came ultimately to an $\mathfrak{L}$-dependent conclusion. Suppose it were in fact the case that the Lagrangian possessed the x-independent structure

$$\mathfrak{L} = \mathfrak{L}(\varphi, \partial\varphi)$$

which the conservation law(s) $\partial_k S^k{}_j = 0$ have been seen to entail. We are in position now to appreciate the importance of the observation that *by gauge transformation, structural features of the Lagrangian—whence also symmetry properties of the associated action functional—can be profoundly altered*. The Lagrangian

$$\mathfrak{L}' = \mathfrak{L}(\varphi, \partial\varphi) + \partial_k \mathcal{G}^k(\varphi, x) = \mathfrak{L}'(\varphi, \partial\varphi, x)$$

will, in general, *not* be x-independent; it serves equally well to describe the physical system in hand (it gives rise to the same field equations), but leads *via* (34) to an $S'^k{}_j$ which is distinct from $S^k{}_j$ and which is, in general, *not* conserved. Had we adopted $\mathfrak{L}'$ at the outset, we would have obtained $\partial_k S'^k{}_j \neq 0$, and would—though they remain valid properties of the system—have *missed* the conservation laws $\partial_k S^k{}_j = 0$. We would have picked up the latter information only if we had thought to ask

Can the x-dependence of $\mathfrak{L}'(\varphi, \partial\varphi, x)$ be "gauged away"?

As was observed already at (23)

$$\mathfrak{L} \longrightarrow \mathfrak{L}' = \mathfrak{L} + \partial_k \mathcal{G}^k \quad \text{induces} \quad S_{\mathcal{R}} \longrightarrow S'_{\mathcal{R}} = S_{\mathcal{R}} + \int_{\partial\mathcal{R}} \mathcal{G}^k d\sigma_k \tag{35}$$

It is the boundary term which, though invisible to Hamilton's Principle, can do violence to applications of Noether's Theorem. Reading from (29), we obtain

$$J^k_r \longrightarrow J'^k_r = J^k_r + \underbrace{\left\{ \left[\Phi_{\alpha r} - \varphi_{\alpha,i} \mathcal{X}^i_r\right] \frac{\partial}{\partial \varphi_{\alpha,k}} + \mathcal{X}^k_r \right\} \left(\partial_j \mathcal{G}^j\right)}_{G^k_r(\varphi, \partial\varphi, x)} \tag{36}$$

In exceptional cases it will be possible to write

$$G^k_r = \partial_j A^{jk}_r \quad \text{with} \quad A^{jk}_r = -A^{kj}_r$$

In such cases—some of which are, as will emerge, physically quite important—the symmetry of the action functional is preserved; one has

$$\partial_k J^k_r = 0 \quad \Longleftrightarrow \quad \partial_k J'^k_r = 0$$

It becomes natural in this light to anticipate that there will arise cases in which it is appropriate to absorb "parameterized gauge transformations" into an enlarged conception of what we are to mean by a "parameterized map," writing

$$\mathcal{T}_{\delta\omega} \quad : \quad \left\{ \begin{array}{rcl} x^i & \longrightarrow & x^i + \delta\omega^r \cdot \mathcal{X}_r^i \\ \varphi_\alpha(x) & \longrightarrow & \varphi_\alpha(x) + \delta\omega^r \cdot \Phi_{\alpha r} \\ \mathcal{L} & \longrightarrow & \mathcal{L} + \delta\omega^r \cdot \partial_j \mathcal{G}_r^j \end{array} \right. \tag{37}$$

in place of (25). Slight adjustment of the argument that gave (29) then gives

$$J_r^k = \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,k}} \Big\{ \Phi_{\alpha r} - \varphi_{\alpha,i} \mathcal{X}_r^i \Big\} + \mathcal{L}\mathcal{X}_r^k + \mathcal{G}_r^k \tag{38}$$

We recall in this connection that in particle mechanics it is precisely such a generalization that makes it possible to construct a Noetherian account of the implications of Galilean covariance.[18]

Each of the statements (32) provides what is, in effect, the differential formulation of a *local conservation law*. What are the associated "conserved quantities?" The question is best approached by looking to the corresponding integral statements

$$\int_{\partial\mathcal{R}} J_r^k d\sigma_k = 0 \quad \text{for all } \mathcal{R}, \text{ with } r = 1, 2, \ldots, \nu \tag{39}$$

Take $\mathcal{R}$ to have, in particular, the form of a "spacetime drum," as illustrated in the figure at the top of the next page. We then have

$$\int_{\text{top}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} + \int_{\text{sides}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} + \int_{\text{bottom}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} = 0$$

The middle term will be assumed to vanish, either because we have imposed spatial boundary conditions of the form $\boldsymbol{J}_r(\text{sides}) = \boldsymbol{0}$ or because we have "pushed the sides of the drum to infinity," where $\boldsymbol{J}_r$ has been assumed to die a natural asymptotic death. The surface differentials $\boldsymbol{d\sigma}$ are, by stipulation of the divergence theorem, all "outer-directed," which on the bottom of the drum means "past-directed." Let us, however, adopt the convention that *surface differentials associated with "timeslices" (surfaces of constant t) will in all cases be "future-directed."* We then have

$$\int_{\text{top}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} + 0 - \int_{\text{bottom}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} = 0$$

or again

$$\int_{\text{top}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} = \int_{\text{bottom}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma}$$

—quite irrespectively of the particular t-value used to position the top of the

[18] See, for example, CLASSICAL MECHANICS (1983), p. 169 and the discussion which appears on pp. 161–170 of CLASSICAL FIELD THEORY (1979).

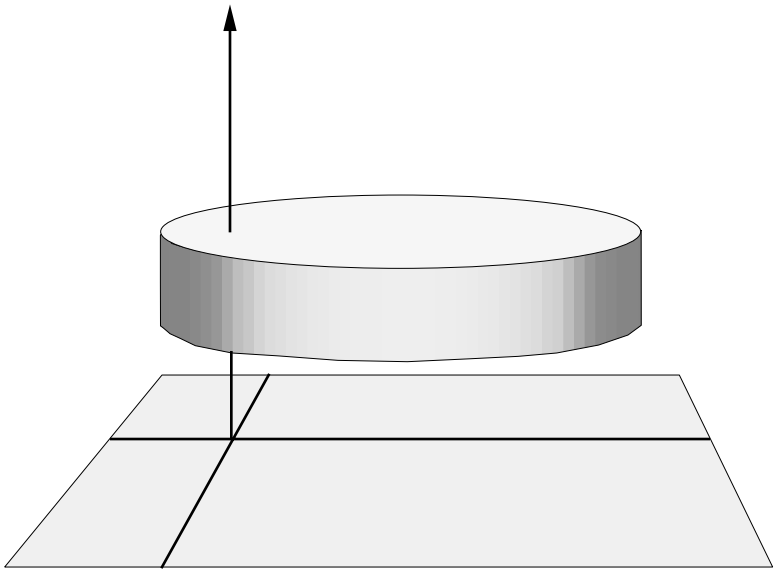


FIGURE 6: *A "drum" in spacetime—a bubble bounded above and below by "timeslices." All points on the top surface have time coordinate t, and all points on the bottom have time coordinate t_0. The edges of the drum may, at the end of the argument, recede to infinity.*

drum. The implication is that the integrated expressions

$$\int_{\text{top}} \boldsymbol{J}_r \cdot \boldsymbol{d\sigma} \equiv \iint \cdots \int J_r^0 dx^1 dx^2 \cdots dx^n \quad : \quad r = 1, 2, \ldots, \nu \tag{40}$$

are (global) *constants of the field motion*.

Returning, by way of illustration, to the translational map (33), we learn from (34) that

$$\left.\begin{aligned} S^0{}_0 &= \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,0}} \varphi_{\alpha,0} - \mathcal{L} \equiv \mathcal{E} \\ S^0{}_1 &= \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,0}} \varphi_{\alpha,1} \qquad \equiv \mathcal{P}_1 \\ S^0{}_2 &= \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,0}} \varphi_{\alpha,2} \qquad \equiv \mathcal{P}_2 \\ S^0{}_3 &= \frac{\partial \mathcal{L}}{\partial \varphi_{\alpha,0}} \varphi_{\alpha,3} \qquad \equiv \mathcal{P}_3 \end{aligned}\right\} \tag{41}$$

Here I have, in the interests of physical concreteness, set $n = 3$; I have assigned x^0 the meaning $x^0 \equiv t$ and have understood $\{x^1, x^2, x^3\}$ to refer to an inertial Cartesian frame in physical 3-space. We note that $\mathcal{E}$ is co-dimensional with $\mathcal{L}$, and has the dimensionality therefore of an *energy density* (energy/volume), while $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$ each has the dimensionality (energy density/velocity)

of *linear momentum density*. We are in position now to assert that if the Lagrangian density has no explicit x-dependence (i.e., is invariant with respect to translations in spacetime), then the following number-valued expressions are global constants of the field motion:

$$E = \int\int\int \mathcal{E}\ dx^1dx^2dx^3 = \text{total energy}$$
$$P_1 = \int\int\int \mathcal{P}_1\, dx^1dx^2dx^3 = \text{total 1-component of linear momentum}$$
$$P_2 = \int\int\int \mathcal{P}_2\, dx^1dx^2dx^3 = \text{total 2-component of linear momentum}$$
$$P_3 = \int\int\int \mathcal{P}_3\, dx^1dx^2dx^3 = \text{total 3-component of linear momentum}$$

The local equations $\partial_k J^k_r = 0$ can now be rendered

$$\tfrac{\partial}{\partial t}(\text{energy density } \mathcal{E}) + \boldsymbol{\nabla}\cdot(\text{energy } \mathbf{flux}) = 0$$
$$\tfrac{\partial}{\partial t}(\text{momentum density } \mathcal{P}_1) + \boldsymbol{\nabla}\cdot(\text{associated momentum } \mathbf{flux}) = 0$$
$$\tfrac{\partial}{\partial t}(\text{momentum density } \mathcal{P}_2) + \boldsymbol{\nabla}\cdot(\text{associated momentum } \mathbf{flux}) = 0$$
$$\tfrac{\partial}{\partial t}(\text{momentum density } \mathcal{P}_3) + \boldsymbol{\nabla}\cdot(\text{associated momentum } \mathbf{flux}) = 0$$

and by straightforward adjustment of the arguments that gave (41) we can obtain *explicit descriptions* of the fluxes in question. These statements assign explicitly detailed meaning to the statement that field energy and momentum, when globally conserved, are conserved because they slosh about in a locally conservative way. And—to restate a point already made—the total energy and momentum of a field system are of manifest "interest" even when they happen *not* to be conserved!

One final remark: at no point in our work thus far have we made actual use of the "group structure" which has been presumed to attach to the infinitesimal parameterized maps which are themselves clearly central to Noether's line of argument. The families of transformations which lay natural claim to our attention do tend generally—spontaneously—to possess the group property, but nowhere have we had to draw upon any of the rich consequences of that fact. That situation will change when we look to details pursuant to certain (important) specific *applications* of Noether's Theorem.

Field-theoretic analog of the Helmholtz conditions. Generalized forces $F_i(q)$ which are "conservative" in the sense they can be derived from a potential

$$F_i(q) = -\frac{\partial}{\partial q^i}U(q)$$

have (owing to the general *equality of the cross derivatives* of $U(q)$) necessarily the property that

$$\frac{\partial F_i}{\partial q^j} - \frac{\partial F_j}{\partial q^i} = 0$$

Conversely (by a famously more difficult line of argument), if $F_i(q)$ possesses the latter property then there exists such a function $U(q)$; it is in fact the case that—in particular consequence of a very general formula due to Poincaré[19]—$U(q)$ admits of this little-known but wonderful explicit construction

$$U(q) = -\int_0^1 F_k(\tau q) q^k d\tau + \text{constant} \tag{42}$$

The preceeding remarks serve to generalize (very slightly) the familiar statement

$$\boldsymbol{A} = \text{grad}\varphi \quad \Longleftrightarrow \quad \text{curl}\boldsymbol{A} = \boldsymbol{0}$$

Identical ideas enter into the observation[20] that the dynamical system

$$\begin{aligned} \dot{q} &= f(q,p) \\ \dot{p} &= g(q,p) \end{aligned}$$

will admit of Hamiltonian formulation if and only if it is true of the functions $f(q,p)$ and $g(q,p)$ that

$$\frac{\partial f}{\partial q} + \frac{\partial g}{\partial p} = 0$$

It seems entirely natural, in the light of such remarks, to ask a question which—quite unaccountably to me—appears in fact to be only very seldom asked: Under what conditions do the coupled second-order differential equations

$$\begin{aligned} G_1(\ddot{q}, \dot{q}, q, t) &= 0 \\ G_2(\ddot{q}, \dot{q}, q, t) &= 0 \\ &\vdots \\ G_n(\ddot{q}, \dot{q}, q, t) &= 0 \end{aligned}$$

admit of Lagrangian formulation

$$G_i(\ddot{q}, \dot{q}, q, t) = \left\{ \frac{d}{dt}\frac{\partial}{\partial \dot{q}^i} - \frac{\partial}{\partial q^i} \right\} L(\dot{q}, q, t)$$

The question was first explored by Hermann von Helmholtz (1821– 1894), who in 1867 established the necessity of the conditions

$$\left.\begin{aligned} \frac{\partial G_i}{\partial \ddot{q}^j} - \frac{\partial G_j}{\partial \ddot{q}^i} &= 0 \\ \frac{\partial G_i}{\partial \dot{q}^j} + \frac{\partial G_j}{\partial \dot{q}^i} &= \frac{d}{dt}\left[\frac{\partial G_i}{\partial \ddot{q}^j} + \frac{\partial G_j}{\partial \ddot{q}^i}\right] \\ \frac{\partial G_i}{\partial q^j} - \frac{\partial G_j}{\partial q^i} &= \frac{1}{2}\frac{d}{dt}\left[\frac{\partial G_i}{\partial \dot{q}^j} - \frac{\partial G_j}{\partial \dot{q}^i}\right] \end{aligned}\right\} \tag{43}$$

[19] See Electrodynamics (1972), p. 173, or p. 14 of my "Electrodynamical Applications of the Exterior Calculus" (1996).

[20] See Classical Mechanics (1983), p. 209.

for which A. Mayer in 1896 established the sufficiency. To establish necessity one has simply to notice that differential equations derived from a Lagrangian are differential equations

$$\begin{aligned} G_i(\ddot{q},\dot{q},q,t) &= \frac{\partial^2 L}{\partial\dot{q}^i\partial\dot{q}^j}\ddot{q}^j + \frac{\partial^2 L}{\partial\dot{q}^i\partial q^j}\dot{q}^j + \frac{\partial^2 L}{\partial\dot{q}^i\partial t} - \frac{\partial L}{\partial q^i} \\ &= g_{ij}(\dot{q},q,t)\ddot{q}^j + h_i(\dot{q},q,t) \end{aligned}$$

of a very particular structure; they are, for example, linear in the 2nd derivatives $\ddot{q}^j$, and the coefficients g_{ij} are necessarily symmetric. The Helmholtz conditions (43) emerge quite naturally when such observations are collected and—this is the point—formulated in such a way as to make no explicit reference to the (generally unknown) Lagrangian itself.[21] In (43) we are presented with an antisymmetric array + a symmetric array + another antisymmetric array of conditions—conditions which in number total

$$\tfrac{1}{2}(n-1)n + \tfrac{1}{2}n(n+1) + \tfrac{1}{2}(n-1)n = \tfrac{1}{2}n(3n-1) = 1,5,12,22,35,\ldots \sim \tfrac{3}{2}n^2$$

The practical utility of the Helmholtz conditions is limited however not so much by their number as by the fact that we never know whether equations that fail the test might by appropriate "pre-processing" be made to pass it; by the converse, that is to say, of the following observation: if equations $G_i = 0$ pass the test, then reorderings, multiplication by factors, formation of linear combinations, etc. will result generally in *equivalent* equations that nevertheless *fail* the test. This awkward circumstance has motivated P. Havas[22] to pose and resolve this more general question: Given an ordered system of equations

$$G_i(\ddot{q},\dot{q},q,t) = 0$$

when do there exist integrating factors $f_i(\dot{q},q,t)$ such that the equivalent system

$$\tilde{G}_i(\ddot{q},\dot{q},q,t) \equiv f_i(\dot{q},q,t)\cdot G_i(\ddot{q},\dot{q},q,t) = 0$$

admit of Lagrangian formulation? Unsurprisingly, the conditions achieved by Havas are markedly more complicated than the Helmholtz conditions. And the Havas conditions, for all their complexity, contribute nothing toward the resolution either of the ordering problem or of the linear combination problem.

Look, by way of illustration, to the simple system

$$G(\ddot{q},\dot{q},q,t) \equiv \ddot{q} + q = 0$$

Here $n = 1$; there is a single Helmholtz conditon, it reads

$$\frac{\partial G}{\partial\dot{q}} = \frac{d}{dt}\left[\frac{\partial G}{\partial\ddot{q}}\right]$$

[21] For the details see pp. 117–120 of Classical Mechanics (1983).

[22] "The range of application of the Lagrange formalism-I," Nuovo Cimento Supp. **5**, 363 (1957).

and is clearly satisfied; there exists an associated Lagrangian (Helmholtz does not tell us how to find it) and by familiar tinkering we know it to be

$$L(\dot{q}, q, t) = \tfrac{1}{2}\dot{q}^2 - \tfrac{1}{2}q^2 + \text{possible gauge term}$$

Look next to the system

$$G(\ddot{q}, \dot{q}, q, t) \equiv \ddot{q} + k\dot{q} + q = 0$$

The single Helmholtz condition now entails $k = 0$, which is the case already studied. We conclude that if $k \neq 0$ then no Lagrangian exists. Suppose, however, we ask this weaker question: Does there exist an integrating factor $f(t)$ such that the equivalent equation

$$\tilde{G}(\ddot{q}, \dot{q}, q, t) \equiv f(t) \cdot G(\ddot{q}, \dot{q}, q, t) = f \cdot (\ddot{q} + k\dot{q} + q) = 0$$

admits of Lagrangian formulation? The Helmholtz condition is seen now to entail $fk = \dot{f}$. We conclude that if $f(t) = f_0 e^{kt}$ then $\tilde{G}(\ddot{q}, \dot{q}, q, t) = 0$ *does* admit of Lagrangian formulation (therefore of Hamiltonian formulation, therefore even of quantum mechanical formulation!), and by unfamiliar tinkering discover the Lagrangian to be given by

$$L(\dot{q}, q, t) = \tfrac{1}{2} f_0 e^{kt}(\dot{q}^2 - q^2) + \text{possible gauge term}$$

We recover the previous Lagrangian at $k = 0$, provided we set the physically inconsequential prefactor f_0 equal to unity. By exercise of some uncommon self-control, I shall forego discussion of some of the interesting physics that can be extracted from generalizations of this striking result.

One gains the impression that from some sufficiently exhaulted formal standpoint the Helmholtz conditions can be understood as but yet another instance of the familiar "curl condition." The reader who wishes to look more closely into that or other aspects of our present topic might be well-advised to start by looking into the dense pages of R. Santilli's *Foundations of Theoretical Mechanics I: The Inverse Problem in Newtonian Mechanics* (1978).

Returning now to field theory, we find it natural, in light of the preceeding discussion, to ask: Under what conditions do the coupled second-order partial differential equations

$$\begin{aligned} G_1(\partial\partial\varphi, \partial\varphi, \varphi, x) &= 0 \\ G_2(\partial\partial\varphi, \partial\varphi, \varphi, x) &= 0 \\ &\vdots \\ G_N(\partial\partial\varphi, \partial\varphi, \varphi, x) &= 0 \end{aligned}$$

admit of Lagrangian formulation

$$G_\alpha(\partial\partial\varphi, \partial\varphi, \varphi, x) = \left\{ \frac{\partial}{\partial x^i} \frac{\partial}{\partial \varphi_{\alpha,i}} - \frac{\partial}{\partial \varphi_\alpha} \right\} \mathcal{L}(\partial\varphi, \varphi, x)$$

Havas, in the introduction to his paper, remarks that "the general problem was studied in great detail by Königsberger,[23] *who also investigated continuous systems*" [my emphasis], so it seems quite possible that the result I am about to describe was known to L. Königsberger and his readers (who, curiously, seem not to have included E. T. Whittaker among their number) already 78 years before it was worked out by me. In any event... if one proceeds in direct imitation of the line of argument which led to (43) one is led[24] to conditions which can be notated

$$\left.\begin{aligned} \frac{\eth G_\alpha}{\eth \varphi_{\beta,ij}} - \frac{\eth G_\beta}{\eth \varphi_{\alpha,ij}} &= 0 \\ \frac{\partial G_\alpha}{\partial \varphi_{\beta,j}} + \frac{\partial G_\beta}{\partial \varphi_{\alpha,j}} &= \frac{\partial}{\partial x^i}\left[\frac{\eth G_\alpha}{\eth \varphi_{\beta,ij}} + \frac{\eth G_\beta}{\eth \varphi_{\alpha,ij}}\right] \\ \frac{\partial G_\alpha}{\partial \varphi_\beta} - \frac{\partial G_\beta}{\partial \varphi_\alpha} &= \frac{1}{2}\frac{\partial}{\partial x^i}\left[\frac{\partial G_\alpha}{\partial \varphi_{\beta,i}} - \frac{\partial G_\beta}{\partial \varphi_{\alpha,i}}\right] \end{aligned}\right\} \tag{44}$$

and concerning which my first obligation is to clarify the meaning and origin of the fancy derivatives. It is an implication of $\varphi_{\alpha,ij} = \varphi_{\alpha,ji}$ that

$$\varphi_{\alpha,ij} = (1-\lambda)\varphi_{\alpha,ij} + \lambda\varphi_{\alpha,ji} \quad (\text{all } \lambda)$$

and follows therefore that

$$F(\ldots, \varphi_{\alpha,ij}, \ldots, \varphi_{\alpha,ji}, \ldots)$$

and its substitutional transform

$$\begin{aligned} &\tilde{F}(\ldots, \varphi_{\alpha,ij}, \ldots, \varphi_{\alpha,ji}, \ldots) \\ &\quad \equiv F(\ldots, (1-\lambda)\varphi_{\alpha,ij} + \lambda\varphi_{\alpha,ji}, \ldots, (1-\mu)\varphi_{\alpha,ji} + \mu\varphi_{\alpha,ij}, \ldots) \end{aligned}$$

are in all cases equal, but (in most cases) functionally distinct. The definition

$$\frac{\eth}{\eth \varphi_{\alpha,ij}} \equiv \frac{1}{2}\left[\frac{\partial}{\partial \varphi_{\alpha,ij}} + \frac{\partial}{\partial \varphi_{\alpha,ji}}\right]$$

has been cooked up to achieve (in all cases, even—trivially—in the case $i = j$)

$$\frac{\eth}{\eth \varphi_{\alpha,ij}} F = \frac{\eth}{\eth \varphi_{\alpha,ij}} \tilde{F}$$

and thus to yield results which are invariant with respect to the exercise of our substitutional options. If N signifies the number of field components, and

[23] *Die Principien der Mechanik* (1901).

[24] See Classical Field Theory (1979), p. 121–124.

$m = n + 1$ the number of independent variables, then by delicate counting we find the conditons (44) to be

$$\begin{aligned}\tfrac{1}{2}(N-1)N \cdot \tfrac{1}{2}m(m+1) + \tfrac{1}{2}N(N+1)\cdot m + \tfrac{1}{2}(N-1)N \\ = \tfrac{1}{4}N\big[N(m^2+3m+2) - (m^2-m+2)\big]\end{aligned}$$

in number. We therefore have

$$\begin{array}{lll}\tfrac{1}{2}N(\ 3N-\ 1) & \text{conditions if} & m=1\\ \tfrac{1}{2}N(\ 6N-\ 2) & \text{conditions if} & m=2\\ \tfrac{1}{2}N(10N-\ 4) & \text{conditions if} & m=3\\ \tfrac{1}{2}N(15N-\ 7) & \text{conditions if} & m=4\\ \tfrac{1}{2}N(21N-11) & \text{conditions if} & m=5\end{array}$$

and recover precisely the Helmholtz conditions (43) in the case $m = 1$. The conditions (44) are beset with all the limitations which have previously been seen to afflict the practical application of the Helmholtz conditions (and are susceptible, I suppose, to the same modes of potential remedy). They are necessary by demonstration, but concerning their sufficiency one can, at this point, only speculate.

Look, by way of illustration, to the class of simple systems

$$G(\partial\partial\varphi, \partial\varphi, \varphi, x) = \varphi_{tt} - \varphi_{xx} + k\varphi^p = 0$$

Here $N = 1$ and $m = 2$; there are two conditions (44), and they read

$$\begin{aligned}\frac{\partial G}{\partial \varphi_t} &= \frac{\partial}{\partial t}\left[\frac{\eth G}{\eth \varphi_{tt}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G}{\eth \varphi_{xt}}\right]\\ \frac{\partial G}{\partial \varphi_x} &= \frac{\partial}{\partial t}\left[\frac{\eth G}{\eth \varphi_{tx}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G}{\eth \varphi_{xx}}\right]\end{aligned}$$

Both are readily seen to be satisfied, so we are entitled to hope (yet cannot, in the absence of a sufficienty proof, be certain) that the system admits of Lagrangian formulation. A little tinkering yields

$$\mathcal{L}(\partial\varphi, \varphi) = \tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\varphi_x^2 - \tfrac{1}{p+1}k\varphi^{p+1}$$

At $k = 0$ we recover the essentials (compare (15)) of the familar wave equation, and at $p = 1$ we obtain the Lagrangian system

$$\varphi_{tt} - \varphi_{xx} + k\varphi = 0$$

which will acquire importance for us as the "Klein-Gordon equation." By easy extension we have

$$\mathcal{L}(\partial\varphi, \varphi) = \tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\varphi_x^2 - F(\varphi) \quad \Longleftrightarrow \quad \varphi_{tt} - \varphi_{xx} + F'(\varphi) = 0$$

In the particular case $F(\varphi) = -k\cos\varphi$ we obtain a much-studied nonlinear field equation

$$\varphi_{tt} - \varphi_{xx} + k\sin\varphi = 0$$

known as the "Sine-Gordon equation," which in the weak-field approximation gives back the (linear) Klein-Gordon equation.

Look next to the system

$$G(\partial\partial\varphi, \partial\varphi, \varphi, x) = \varphi_t - \varphi_{xx} = 0$$

which captures the essence of the so-called "heat equation" (otherwise known as the "diffusion equation"). From the same pair of conditions as served us in the preceeding example we are led promptly to the conclusion that the heat equation does *not* admit of Lagrangian formulation. Nor does the obvious variant of the "integrating factor trick" salvage the situation. That I would encourage my reader to regard as a profoundly unsatisfactory state of affairs, as an invitation to invention. For the heat equation is an important thing, and the Lagrangian formalism is an important thing, and it is "unreasonable" that they should have nothing to say to each other. Since we will fairly frequently find ourselves in analogous predicaments, I digress to illustrate the kind of escape routes that can, with a little cleverness, be devised. One standard trick—the "auxiliary field trick," as it is sometimes called—hinges on the recognition that our field may possess a heretofore overlooked "companion field." With such an idea in mind, it does in the present context not take one long to concoct the 2-field system

$$\mathcal{L} = \tfrac{1}{2}(\alpha_t\varphi - \alpha\varphi_t) - \alpha_x\varphi_x$$

and to observe that

$$\left\{\frac{\partial}{\partial t}\frac{\partial}{\partial \alpha_t} + \frac{\partial}{\partial x}\frac{\partial}{\partial \alpha_x} - \frac{\partial}{\partial \alpha}\right\}\mathcal{L} = 0 \quad \text{gives} \quad \varphi_{xx} = +\varphi_t$$

$$\left\{\frac{\partial}{\partial t}\frac{\partial}{\partial \varphi_t} + \frac{\partial}{\partial x}\frac{\partial}{\partial \varphi_x} - \frac{\partial}{\partial \varphi}\right\}\mathcal{L} = 0 \quad \text{gives} \quad \alpha_{xx} = -\alpha_t$$

We have successfully reproduced the diffusion equation, but at the cost of introducing an auxiliary field that satisfies an "antidiffusion equation," and the question becomes "What are we to make of that?" Have we made a discovery, or merely replaced one embarrassment by another? Such questions tend generally to be important questions, but to admit of no easy (or at least of no universal) answer; everything hinges on details of the particular system at hand.[25] The natural force of well-crafted formalism is frequently more persuasive—and a surer guide—than the imperfect evidence of a laboratory. But tact and good

[25] Later I will have occasion to develop the sense in which manipulations quite similar to those just sketched lead with a kind of inevitability to "the invention of quantum mechanics!" One would appear to be well within one's rights to say of such a development that it is "important," and has the character of a "discovery."

judgment are all-important, for the free-spinning formalist teeters always on the edge of triviality. For example:

Allegations of the form $A = B$ can—whether true or false—always be displayed as "variational principles"

$$(A - B)^2 = \text{minimum}$$

but it would be frivilous to claim that the elegance of the display enhances the likelihood that the allegation is correct.[26] Somewhat less frivilous is the observation that, while (as we have seen) the system

$$\begin{aligned}\dot{q} &= f(q,p)\\ \dot{p} &= g(q,p)\end{aligned}$$

admits of Hamiltonian formulation only when a certain "curl condition" is satisfied, it can always be absorbed into an expanded system which does admit of such formulation: construe q and p to be "coordinates in a 2-space," let Q and P denote their conjugate momenta, and construct

$$H(q,Q,p,P) = Qf(q,p) + Pg(q,p)$$

Then

$$\begin{aligned}\dot{q} &= +\partial H/\partial Q = f(q,p)\\ \dot{p} &= +\partial H/\partial P = g(q,p)\end{aligned}$$

which have, however, acquired these companions:

$$\begin{aligned}\dot{Q} &= -\partial H/\partial q = -Q\frac{\partial f}{\partial q} - P\frac{\partial g}{\partial q}\\ \dot{P} &= -\partial H/\partial p = -Q\frac{\partial f}{\partial p} - P\frac{\partial g}{\partial p}\end{aligned}$$

It's an admittedly "cheap trick," but on occasion proves useful. Similarly "cheap" is the observation that if we wish to achieve

$$G(\partial\partial\varphi, \partial\varphi, \varphi, x) = 0$$

and are not opposed to the free introduction of auxiliary fields, then to achieve contact with the Lagrangian formalism we have only to constuct

$$\mathcal{L}(\partial\partial\varphi, \partial\varphi, \varphi, \alpha, x) = -\alpha G(\partial\partial\varphi, \partial\varphi, \varphi, x)$$

Trivially,

$$\left\{\frac{\partial}{\partial x^i}\frac{\partial}{\partial\left(\frac{\partial\alpha}{\partial x^i}\right)} - \frac{\partial}{\partial\alpha}\right\}\mathcal{L} = G(\partial\partial\varphi, \partial\varphi, \varphi, x) = 0$$

[26] See, however, the discussion of Gauß' "Principle of Least Constraint" which can be found in §105 of Whittaker's *Analytical Mechanics*.

but we acquire an obligation to describe also the motion of the auxiliary field, and that (since the Lagrangian depends now—non-standardly—also upon the *second* partials of φ) requires that we subject the Lagrangian formalism to a bit of stretching, along lines first explored in (1850) by M. Ostrogradsky.[27] Thus are we led at length to write (if I may lapse for a moment into the notation appropriate to the descripton of a *multi-component* field system $\varphi = \{\varphi_\alpha\}$)

$$\left\{ -\frac{\partial^2}{\partial x^i \partial x^j}\frac{\partial}{\partial \varphi_{\alpha,ij}} + \frac{\partial}{\partial x^i}\frac{\partial}{\partial \varphi_{\alpha,i}} - \frac{\partial}{\partial \varphi_\alpha} \right\} \mathfrak{L} = 0 \tag{45}$$

where terms containing $\varphi_{\alpha,ijk}$ will be avoided if and only if $\varphi_{\alpha,ij}$ enters at most linearly into the structure of the functions $G(\partial\partial\varphi, \partial\varphi, \varphi, x)$. Returning again to our most recent example, we construct

$$\mathfrak{L}' = -\alpha(\varphi_t - \varphi_{xx})$$

and, in addition to the diffusion equation, recover (by application of (45)) precisely the backwards diffusion equation $\alpha_t + \alpha_{xx} = 0$. Nor is this, in fact, surprising, for

$$\begin{aligned} \mathfrak{L} - \mathfrak{L}' &= \tfrac{1}{2}\alpha_t\varphi + \tfrac{1}{2}\alpha\varphi_t - \alpha_x\varphi_x - \alpha\varphi_{xx} \\ &= (\tfrac{1}{2}\alpha_t\varphi)_t + (-\alpha\varphi_x)_x \end{aligned}$$

shows that $\mathfrak{L}$ and $\mathfrak{L}'$ are in fact *gauge equivalent*. Here a "seeming triviality" and a "stroke of modest genius" are seen to be actually of identical force. One should perhaps not be too casually dismissive of trivialities.

Hamiltonian methods in classical field theory. We wrote out the Lagrangian theory of a one-dimensional crystal, traced that theory to its continuous limit where we obtained a Lagrangian description of the dynamics of a string, and by straightforward generalization we led to an elegantly functional "Lagrangian formulation of the classical dynamics of field systems." Plausibly that same strategy would lead us to a "Hamiltonian fomulation of classical field theory," to the associated "theory of canonical transformations," whence finally to a "field-theoretic generalization of Hamilton-Jacobi theory." Alternatively, we might attempt to work entirely within a field-theoretic framework—building upon the Lagrangian formalism in direct imitation of particle-theoretic practice, but avoiding all reference to "crystals," to "refinement of the lattice." We might expect to be led ultimately—by natural extension of standard quantization procedures—to a "quantum theory of fields."

Central to all such formal elaboration, we can anticipate, will be a field-theoretic analog of the "conjugate momentum" concept. Were we to proceed by the lattice-refinement technique, we can expect to achieve[28]

$$p_i = p_i(\dot{q}, q, t) \equiv \frac{\partial L}{\partial \dot{q}^i} \xrightarrow[\text{lattice refinement}]{} \pi_\alpha = \pi_\alpha(\partial\varphi, \varphi, x) \equiv \frac{\partial \mathfrak{L}}{\partial \dot{\varphi}^\alpha}$$

[27] See Whittaker's §110.

[28] At this point it becomes formally more natural to write φ^α where formerly we have written φ_α.

And there's the rub. Field theory presents us with a *population* of co-equal constructions

$$\frac{\partial \mathfrak{L}}{\partial\left(\frac{\partial \varphi^\alpha}{\partial t}\right)},\ \frac{\partial \mathfrak{L}}{\partial\left(\frac{\partial \varphi^\alpha}{\partial x^1}\right)},\ \frac{\partial \mathfrak{L}}{\partial\left(\frac{\partial \varphi^\alpha}{\partial x^2}\right)},\ \cdots,\ \frac{\partial \mathfrak{L}}{\partial\left(\frac{\partial \varphi^\alpha}{\partial x^n}\right)}$$

and it would appear to be in violation of the essential spirit of field theory (and certainly in violation of the spirit of relativity) to promote one of those—equivalently, to promote one of the independent variables $\{t, x^1, x^2, \ldots, x^n\}$—to a status of preeminence over the others. What to do? Should we associate with each field component an indexed *set* of "conjugate momenta"

$$\varphi^\alpha \quad \longleftrightarrow \quad \left\{ \begin{array}{l} \pi_{\alpha 0} = \frac{\partial \mathfrak{L}}{\partial(\partial \varphi^\alpha / \partial x^0)} \\ \pi_{\alpha 1} = \frac{\partial \mathfrak{L}}{\partial(\partial \varphi^\alpha / \partial x^1)} \\ \qquad \vdots \\ \pi_{\alpha n} = \frac{\partial \mathfrak{L}}{\partial(\partial \varphi^\alpha / \partial x^n)} \end{array} \right.$$

Such a procedure would appear to do such radical violence to the essentials of Hamiltonian mechanics as to be unworkable. Should the $\pi_{\alpha i}$ be made to participate co-equally in the assembly into some unitary object along (perhaps) these general lines

$$\pi_\alpha = \sum_{i=0}^{n} \pi_{\alpha i} P^i$$

To do so would be to import into the theory an auxiliary object P^i which is unprecidented in the Hamiltonian mechanics of particles, and of which no natural candidate presents itself.

By way of preparaion for what follows, I digress now to observe that one can bring geometrical imagery to $\phi(x, y, z)$ (which I shall, in service of concreteness, assume to be number-valued and real) in a variety of distinct ways; one can consider that $\phi(x, y, z)$ describes

a "point" in an (∞-dimensional) space of functions $f(x, y, z)$

a x-parameterized "curve" in a space of functions $f(y, z)$
a y-parameterized "curve" in a space of functions $f(x, z)$
a z-parameterized "curve" in a space of functions $f(x, y)$

a (x, y)-parameterized "curve" in a space of functions $f(z)$
a (x, z)-parameterized "curve" in a space of functions $f(y)$
a (y, z)-parameterized "curve" in a space of functions $f(x)$

a (x, y, z)-parameterized point on the real line

Lagrangian field theory proceeds implicitly from the first point of view, though when we, as physicists, undertake to comprehend what $\varphi(t, x, y)$ is telling us (or to discover the implications of prescribed initial data) we frequently have recourse to the second viewpoint: we make mental "movies." The Hamiltonian formalism(s) developed below embrace the second viewpoint explicitly, from the outset.

Let $\{x^0, x^1, \ldots, x^n\}$ refer to some specified coordinatization of spacetime (of which there are, of course, infinitely many). Identically structured and physically equivalent Hamiltonian formalisms $\mathcal{H}^i$ of $(n+1)$ distinct "flavors" come into existence as follows: In $\mathcal{H}^i$ the variable x^i has been promoted to a distinguished status; it rules as "the parameter." It is the business of $\mathcal{H}^i$ to inscribe "x^i-parameterized dynamical flow curves" in the functional analog Γ^i of a $2N$-dimensional phase space, as illustrated in the following figure. Each

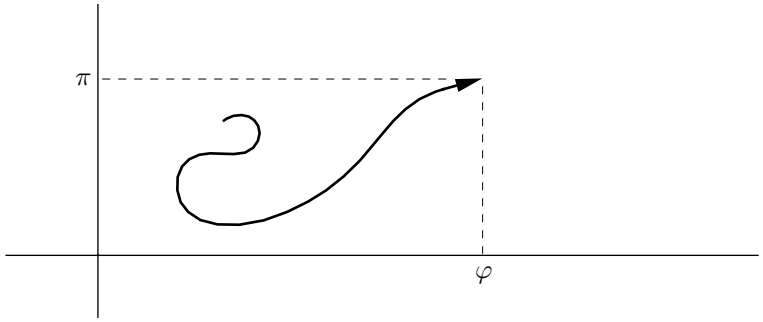


FIGURE 7: *The figure provides a highly schematic representation of a* "dynamical flow curve in the phase space Γ^i." *In the figure,* φ *refers to the* N*-tuple of field functions* $\varphi^\alpha(x)$*, construed to be functions of the variables*

$$x^0, x^1, x^2, \ldots, \bullet, \ldots, x^n \quad \textit{(the missing variable is } x^i\textit{)}$$

and of the parameter x^i*, while* π *refers to the conjugate fields*

$$\pi_\alpha(x) \equiv \frac{\partial \mathcal{L}}{\partial(\partial\varphi^\alpha/\partial x^i)}$$

—similarly construed.

individual $\mathcal{H}^i$ proceeds from a symmetry-breaking act—the promotion of an arbitrarily selected variable to "distinguished status." But construction of the composit formalism

$$\mathcal{H} = \mathcal{H}^0 \oplus \mathcal{H}^1 \oplus \cdots \oplus \mathcal{H}^n$$

entails *no* such symmetry-breaking act. It becomes attractive in this light to regard the field-theoretic "Hamiltonian formalism" to reside in $\mathcal{H}$; to do so would be to embrace a formalism with many redundant component parts, and to accept an obligation ultimately to discuss the (transformation-theoretic and other) interconnections amongst those parts.

In the discussion that follows I shall, in place of x^i, write t to denote the promoted variable. The parameter t can—and in the literature[29] typically

[29] See, for example, H. Goldstein, *Classical Mechanics* (2nd edition 1980), §12–4.

does—refer physically to "time," but should in the generic case be taken to refer to some word like "typical." The generic instance of Hamiltonian field theory will be denoted $\mathcal{H}^t$. Field derivatives with respect to the parameter t will be written $\dot{\varphi}$, while derivatives with respect to the independent variables $\{x^1, x^2, \ldots, x^n\}$ will collectively be denoted ∇x. Where formerly we were content to write $\mathcal{L}(\partial\varphi, \varphi, x)$ we would, by these more explicit conventions, write $\mathcal{L}(\dot{\varphi}, \nabla\varphi; \varphi; t, x)$. We will soon come to attach interest to the observation that the field equations themselves, by these conventions, admit of the following curious reformulation:

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{\varphi}^\alpha} - \left\{ \frac{\partial \mathcal{L}}{\partial \varphi^\alpha} - \frac{\partial}{\partial x^k} \sum_{k-1}^{n} \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \varphi^\alpha}{\partial x^k}\right)} \right\} = 0$$

Equivalently (since $\mathcal{L}$ has been assumed not to depend upon second derivatives of the fields, and is in particular therefore $\nabla\dot{\varphi}$-independent)

$$\frac{\partial}{\partial t}\left\{ \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^\alpha} - \frac{\partial}{\partial x^k} \sum_{k-1}^{n} \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \dot{\varphi}^\alpha}{\partial x^k}\right)} \right\} - \left\{ \frac{\partial \mathcal{L}}{\partial \varphi^\alpha} - \frac{\partial}{\partial x^k} \sum_{k-1}^{n} \frac{\partial \mathcal{L}}{\partial\left(\frac{\partial \varphi^\alpha}{\partial x^k}\right)} \right\} = 0$$

More compactly

$$\frac{\partial}{\partial t}\frac{\delta \mathcal{L}}{\delta \dot{\varphi}^\alpha} - \frac{\delta \mathcal{L}}{\delta \varphi^\alpha} = 0 \tag{46}$$

where

$$\frac{\delta \mathcal{L}}{\delta \varphi^\alpha} \equiv \left\{ \frac{\partial}{\partial \varphi^\alpha} - \frac{\partial}{\partial x^k} \sum_{k-1}^{n} \frac{\partial}{\partial\left(\frac{\partial \varphi^\alpha}{\partial x^k}\right)} \right\} \mathcal{L} \tag{47}$$

serves to define the so-called "variational derivative" of $\mathcal{L}$ with respect to φ^α. Equations (46), which exist in as many variants as there are ways to assign specific meaning to the parameter t, manage—"by contrivance," as it were—to resemble the Lagrange equations of particle mechanics, and put us in position to write out the generic Hamiltonian field theory $\mathcal{H}^t$ by proceeding in formal mimicry of the methods of particle mechanics.

Our former definition of the momentum fields π_α which are (within $\mathcal{H}^t$) conjugate to the φ^α-fields can now be notated

$$\pi_\alpha = \frac{\delta \mathcal{L}}{\delta \dot{\varphi}^\alpha} \tag{48}$$

Evidently $[\varphi^\alpha][\pi_\alpha] = [\mathcal{L}][\,\textit{physical dimension of the parameter } t\,]$, which when t has the nature of a "time"—but not otherwise—entails

$$[\varphi^\alpha][\pi_\alpha] = \text{ action density} \tag{49}$$

We note also that—whatever the degree of "directly physical significance" that attaches to the φ-fields (and that varies from application to application)—we

should in no case assign "direct physical significance" to the associated π-fields, for they respond non-invariantly to gauge transformations:

$$\mathcal{L} \longrightarrow \mathcal{L} + \text{gauge term} \qquad \text{induces} \qquad \pi_\alpha \longrightarrow \pi_\alpha + \text{gauge term}$$

If the so-called "Hessian" (the determinant of the matrix $\|\partial^2\mathcal{L}/\partial\dot{\varphi}^\alpha\partial\dot{\varphi}^\beta\|$) does not vanish, then it becomes possible-in-principle—by functional inversion of the system of equations

$$\pi_\alpha = \frac{\delta}{\delta\dot{\varphi}^\alpha}\mathcal{L}(\dot{\varphi}, \nabla\varphi; \varphi; t, x) = \pi_\alpha(\dot{\varphi}, \nabla\varphi; \varphi; t, x)$$

—to write $$\dot{\varphi}^\alpha = \dot{\varphi}^\alpha(\pi, \nabla\varphi; \varphi; t, x)$$

and therefore to construct that particular "Legendre transform" of $\mathcal{L}$

$$\begin{aligned}\mathcal{H} &= \pi_\alpha\dot{\varphi}^\alpha - \mathcal{L}(\dot{\varphi}, \nabla\varphi; \varphi; t, x)\Big|_{\dot{\varphi}^\alpha \longrightarrow \dot{\varphi}^\alpha(\pi, \nabla\varphi; \varphi; t, x)} \\ &= \mathcal{H}(\pi, \nabla\varphi; \varphi; t, x)\end{aligned} \tag{50}$$

which (within the $\mathcal{H}^t$ formalism) plays the role of a "Hamiltonian density." The associated "Hamiltonian" would be constructed

$$H = \int \mathcal{H} dx^1 \dots dx^n$$

Clearly, $[\mathcal{H}] = [\mathcal{L}] =$ *energy density* and $[H] =$ *energy*. It is notable that $\mathcal{H}$, as displayed in (50), is devoid of the $\nabla\pi$-dependence one might, on formal grounds, otherwise have expected. So far as concerns its $\nabla\varphi$-dependence, we compute

$$\begin{aligned}\frac{\partial\mathcal{H}}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)} &= \underbrace{\pi_\beta\frac{\partial\dot{\varphi}^\beta}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)} - \frac{\partial\mathcal{L}}{\partial\dot{\varphi}^\beta}\frac{\partial\dot{\varphi}^\beta}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)}}_{=\,0 \text{ by the definition of } \pi_\beta} - \frac{\partial\mathcal{L}}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)} \\ &= -\frac{\partial\mathcal{L}}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)}\end{aligned} \tag{51}$$

of which we will have immediate need. For by a similar calculation

$$\begin{aligned}\frac{\partial\mathcal{H}}{\partial\varphi^\alpha} &= \underbrace{\pi_\beta\frac{\partial\dot{\varphi}^\beta}{\partial\varphi^\alpha} - \frac{\partial\mathcal{L}}{\partial\dot{\varphi}^\beta}\frac{\partial\dot{\varphi}^\beta}{\partial\varphi^\alpha}}_{0} - \frac{\partial\mathcal{L}}{\partial\varphi^\alpha} \\ &= -\frac{\partial}{\partial t}\pi_\alpha - \sum_{k=1}^{n}\frac{\partial}{\partial x^k}\frac{\partial\mathcal{L}}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)} \quad \text{by the field equations}\end{aligned}$$

and when we draw upon (51) we obtain

$$\begin{aligned}\frac{\partial}{\partial t}\pi_\alpha &= -\left\{\frac{\partial \mathcal{H}}{\partial \varphi^\alpha} - \sum_{k=1}^{n}\frac{\partial}{\partial x^k}\frac{\partial \mathcal{H}}{\partial\left(\frac{\partial \varphi^\alpha}{\partial x^k}\right)}\right\} \\ &= -\frac{\delta \mathcal{H}}{\delta \varphi^\alpha}\end{aligned} \tag{52}$$

Finally—by a simpler variant of the same line of argument—we have

$$\frac{\partial \mathcal{H}}{\partial \pi_\alpha} = \dot{\varphi}^\alpha + \underbrace{\pi_\beta \frac{\partial \dot{\varphi}^\beta}{\partial \pi_\alpha} - \frac{\partial \mathcal{L}}{\partial \dot{\varphi}^\beta}\frac{\partial \dot{\varphi}^\beta}{\partial \pi_\alpha}}_{0} = \dot{\varphi}^\alpha$$

which, owing to the fact that (as has already been remarked) $\mathcal{H}$ is actually $\nabla\pi$-dependent, can be notated

$$\frac{\partial}{\partial t}\varphi^\alpha = +\frac{\delta \mathcal{H}}{\delta \pi_\alpha} \tag{53}$$

Pulling these results together, we have

$$\left.\begin{aligned}\dot{\varphi}^\alpha &= +\frac{\delta \mathcal{H}}{\delta \pi_\alpha} \\ \dot{\pi}_\alpha &= -\frac{\delta \mathcal{H}}{\delta \varphi^\alpha}\end{aligned}\right\} \tag{54}$$

which are the field equations in "canonical Hamiltonian form." The derivation of (54) has been designed to resemble maximally its counterpart in particle mechanics.[30] We recognize the cancellations encountered along the way to be a characteristic signature of the Legendre transform, seen in all of its diverse applications.

One might plausibly suppose, in view of the structure of (46) and of (54), that we are off and running; that we are now in position to work our way through (say) Whittaker or Goldstein, painlessly translating the concepts and formulæ basic to the analytical dynamics of particles—one after another, as we come to them—into the language of field theory. For example, we might find it natural in imitation of

$$[A, B] \equiv \sum_{k=1}^{n}\left\{\frac{\partial A}{\partial q^k}\frac{\partial B}{\partial p_k} - \frac{\partial B}{\partial q^k}\frac{\partial A}{\partial p_k}\right\}$$

[30] See, for example, p. 193 of Classical Mechanics (1983). An alternative line of argument, which proceeds not from properties of $\mathcal{H}$ but from properties of the Hamiltonian $H = \int \mathcal{H} dx$ and uses "integration by parts" to motivate the introduction of the variational derivative, see Chapter IV, pp. 19–20 of Classical Theory of Fields (1965).

to introduce a "field-theoretic Poisson bracket" by

$$[\mathcal{A},\mathcal{B}] \equiv \sum_{\varkappa=1}^{N}\left\{\frac{\delta\mathcal{A}}{\delta\varphi^{\varkappa}}\frac{\delta\mathcal{B}}{\delta\pi_{\varkappa}} - \frac{\delta\mathcal{B}}{\delta\varphi^{\varkappa}}\frac{\delta\mathcal{A}}{\delta\pi_{\varkappa}}\right\} \tag{55}$$

and to notice that the canonical equations (54) can in this notation be written

$$\left.\begin{aligned}\dot{\varphi}^{\alpha} &= [\varphi^{\alpha},\mathcal{H}]\\ \dot{\pi}_{\alpha} &= [\pi_{\alpha},\mathcal{H}]\end{aligned}\right\} \tag{56}$$

One appears at this point to possess the seed of a notion of "$\mathcal{H}$-generated flow in phase space"

$$\varphi^{\alpha}(t) \longrightarrow \varphi^{\alpha}(t+dt) = \varphi^{\alpha}(t) + dt\cdot[\varphi^{\alpha},\mathcal{H}]$$

and to stand on the brink of a field-theoretic analog of the theory of canonical transformations. But all is not quite so simple. Surprises lurk... for reasons which have partly to do with the circumstance that $\mathcal{H}(\pi,\nabla\varphi;\varphi;t,x)$ possesses —in its $\nabla\varphi$-dependence—a structural element which is absent from $H(p,q)$. I turn now to a discussion intended to identify more clearly some of the points at issue.

In the Hamiltonian mechanics of particles, the familiar constructions

$$\begin{aligned}\text{coordinate} &= q\\ \text{conjugate momentum} &= p\\ \text{energy} &= \tfrac{1}{2m}p^2 + U(x)\\ \text{angular momentum} &= xp_y - yp_x\\ &\vdots\end{aligned}$$

make it natural to assign the name "observable" to functions of the type

$$A = A(p,q;t) = A(p_1,p_2,\ldots,p_n,q^1,q^2,\ldots,q^n;t)$$

and to notice that, in consequence of the canonical equations of motion,

$$\dot{A} = [A,H] + \frac{\partial}{\partial t}A \tag{57}$$

By natural extension (taking care to cast our net wide enough to include $\mathcal{H}$ itself; see again (50)), we assign the name "observable density" to constructions of the type

$$\mathcal{A} = \mathcal{A}(\nabla\pi,\nabla\varphi,\pi,\varphi,x;t)$$

and look to the evaluation of

$$\dot{\mathcal{A}} = \frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\pi_{\alpha}}{\partial x^{k}}\right)}\frac{\partial}{\partial t}\frac{\partial\pi_{\alpha}}{\partial x^{k}} + \frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\varphi^{\alpha}}{\partial x^{k}}\right)}\frac{\partial}{\partial t}\frac{\partial\varphi^{\alpha}}{\partial x^{k}} + \frac{\partial\mathcal{A}}{\partial\pi_{\alpha}}\dot{\pi}_{\alpha} + \frac{\partial\mathcal{A}}{\partial\varphi^{\alpha}}\dot{\varphi}^{\alpha} + \frac{\partial\mathcal{A}}{\partial t}$$

The trick here is to notice that

$$\begin{aligned}\frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\pi_\alpha}{\partial x^k}\right)}\frac{\partial}{\partial t}\frac{\partial\pi_\alpha}{\partial x^k} &= \frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\pi_\alpha}{\partial x^k}\right)}\frac{\partial}{\partial x^k}\dot{\pi}_\alpha \\ &= \frac{\partial}{\partial x^k}\left\{\frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\pi_\alpha}{\partial x^k}\right)}\dot{\pi}_\alpha\right\} - \dot{\pi}_\alpha\frac{\partial}{\partial x^k}\frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\pi_\alpha}{\partial x^k}\right)}\end{aligned}$$

and that the corresponding φ-term yields to similar manipulation. We therefore have

$$\begin{aligned}\dot{\mathcal{A}} = &\left\{\frac{\partial\mathcal{A}}{\partial\varphi^\alpha} - \frac{\partial}{\partial x^k}\frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)}\right\}\dot{\varphi}^\alpha + \left\{\frac{\partial\mathcal{A}}{\partial\pi_\alpha} - \frac{\partial}{\partial x^k}\frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\pi_\alpha}{\partial x^k}\right)}\right\}\dot{\pi}_\alpha \\ &+ \frac{\partial}{\partial x^k}\left\{\frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)}\dot{\varphi}^\alpha + \frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\pi_\alpha}{\partial x^k}\right)}\dot{\pi}_\alpha\right\} + \frac{\partial\mathcal{A}}{\partial t}\end{aligned}$$

and if we draw upon the canonical equations (54), recall the definition (47) of the variational derivative, and note that

$$\frac{\delta\mathcal{A}}{\delta\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)} = \frac{\partial\mathcal{A}}{\partial\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)} \quad \text{because } \mathcal{A} \text{ is } \nabla\nabla\mathcal{A}\text{-independent}$$

we obtain

$$\begin{aligned}\dot{\mathcal{A}} = &\left\{\frac{\delta\mathcal{A}}{\delta\varphi^\alpha}\frac{\delta\mathcal{H}}{\delta\pi_\alpha} - \frac{\delta\mathcal{H}}{\delta\varphi^\alpha}\frac{\delta\mathcal{A}}{\delta\pi_\alpha}\right\} + \frac{\partial\mathcal{A}}{\partial t} \\ &+ \frac{\partial}{\partial x^k}\left\{\frac{\delta\mathcal{A}}{\delta\left(\frac{\partial\varphi^\alpha}{\partial x^k}\right)}\frac{\delta\mathcal{H}}{\delta\pi_\alpha} - \frac{\delta\mathcal{H}}{\delta\varphi^\alpha}\frac{\delta\mathcal{A}}{\delta\left(\frac{\partial\pi_\alpha}{\partial x^k}\right)}\right\}\end{aligned} \tag{58}$$

At this point field theory and particle mechanics appear to have diverged, for the dangling term in (58)—which is present except in special cases of the type $\mathcal{A}(\pi,\varphi)$—has no counterpart in (57). If, however, we allow ourselves to write

$$\text{“observable”} = \int \text{“observable density”} dx^1\cdots dx^n$$

then we can proceed from (58) to the conclusion

$$\begin{aligned}\dot{A} &= \int\dot{\mathcal{A}}dx \\ &= \int\left\{[\mathcal{A},\mathcal{H}] + \frac{\partial\mathcal{A}}{\partial t}\right\}dx + \text{surface term} \\ &\downarrow \\ &= \int\left\{[\mathcal{A},\mathcal{H}] + \frac{\partial\mathcal{A}}{\partial t}\right\}dx \quad \text{when the “surface term” vanishes}\end{aligned} \tag{59}$$

At (55) we assigned meaning to the “Poisson bracket of a pair of observable *densities*.” The result just achieved invites us to write

$$[A,B] \equiv \int[\mathcal{A},\mathcal{B}]dx \tag{60}$$

and thus to assign meaning to the "Poisson bracket of a pair of *observables*." The two concepts are clearly related, yet clearly distinct. If we return with this notation to (59) we find ourselves writing an equation which is identical in appearance to (57), but which carries now a field-theoretic meaning. In particular, if hypotheses sufficient to force the "surface term" to vanish are in place, and if $\partial\mathcal{A}/\partial t = 0$ (i.e., if the observable A is devoid of any *explicit* t-dependence) then

$$\dot{A} = [A, H] \tag{61}$$

From (61) we can recover (56) as special cases, and are led to the conclusion that

$$[A, H] = 0 \quad \Longrightarrow \quad A = \int \mathcal{A} dx \text{ is a } \textit{constant of the field-motion}$$

Evidently $[A, H] = 0$ provides a global formulation of the local statement $[\mathcal{A}, \mathcal{H}] = 0$.

It becomes instructive at this point to revisit our former discussion of Noether's Theorem, which recent remarks[31] will have recalled to the minds of attentive readers. Noether was led from the specification of certain "maps" to the construction at (29) of certain expressions $J_r^k(\varphi, \partial\varphi, x)$—constructions which within the Hamiltonian formalism $\mathcal{H}^i$ acquire the status of "observable densities" of a particular design:

$$\left.\begin{aligned} \mathcal{J}_r^0 &= \qquad \pi_\alpha \left\{ \Phi_r^\alpha - \big(\dot{\varphi}^\alpha \mathcal{X}_r^0 + \sum \frac{\partial \varphi^\alpha}{\partial x^k} \mathcal{X}_r^k\big) \right\} + \big(\pi_\alpha \dot{\varphi}^\alpha - \mathcal{H}\big)\mathcal{X}_r^0 \\ \mathcal{J}_r^i &= \frac{\partial \mathcal{L}}{\partial\big(\frac{\partial \varphi^\alpha}{\partial x^i}\big)} \left\{ \Phi_r^\alpha - \big(\dot{\varphi}^\alpha \mathcal{X}_r^0 + \sum \frac{\partial \varphi^\alpha}{\partial x^k} \mathcal{X}_r^k\big) \right\} + \big(\pi_\alpha \dot{\varphi}^\alpha - \mathcal{H}\big)\mathcal{X}_r^i \end{aligned}\right\} \tag{62}$$

Here $\mathcal{X}_r^k(t, x)$ and $\Phi_r^\alpha(\varphi; t, x)$ are considered to have been prescribed, and $\dot{\varphi}^\alpha$ is to be read as it was during the assembly at (50) of $\mathcal{H}$; i.e., as a reference to $\dot{\varphi}_r^\alpha(\pi, \nabla\varphi; \varphi; t, x)$. Local conservation laws rooted in Noether's Theorem—rooted, that is to say, in the Lagrangian formalism—have (recall (32)) the characteristic form

$$\partial_k J_r^k = 0$$

Such statements are, as I have already emphasized, to be read as *implications of the field equations*; i.e., of the Lagrange equations of motion, of which they express a revealed symmetry property. But, so far as I am aware, there exists no generally-applicable technique for explicitly *demonstrating* that

$$\text{field equations} \quad \Longrightarrow \quad \partial_k J_r^k = 0$$

Such problems are tackeled case-by-case, by *ad hoc* methods special to the instance, as witnessed in the discussion subsequent to (34). But consider: Noetherian conservation laws admit (see again (40)) of *global* formulation

$$\dot{J}_r = 0 \quad \text{with} \quad J_r = \int J_r^0 dx$$

[31] Compare p. 32.

And this is a statement which translates directly into language natural to the Hamiltonian formalism

$$\dot{J}_r = 0 \quad \text{with} \quad J_r = \int \mathcal{J}^0_r dx$$

where it leads to a local statement

$$[\, \mathcal{J}^0_r, \mathcal{H}\,] = 0$$

which is in every computational respect quite distinct from its Lagrangian counterpart, $\partial_k J^k_r = 0$. Interestingly, what was a bothersome gap within the Lagrangian formalism is a question rendered moot in the Hamiltonian formalism; a generally-applicable technique for explicitly demonstrating *how it comes about* that

$$\text{canonical field equations} \quad \Longrightarrow \quad [\, \mathcal{J}^0_r, \mathcal{H}\,] = 0$$

is written onto the very face of the statement. Interesting also is the fact that, while statements of the form $\partial_k J^k = 0$ just sit there as unitary thoughts within the Lagrangian formalism, they admit of as many (generally distinct and complementary) modes of Hamiltonian interpretiation as there are variants of the Hamiltonian formalism. I draw attention finally to the fact that the densities $\mathcal{J}^0_k(\nabla\varphi, \varphi, x; t)$ described by (62) are structurally quite particular; evidently one cannot expect to construct, in any natural way, a "Noetherian interpretation" of the conservation law $[\mathcal{A}, \mathcal{H}\,] = 0$ if $\mathcal{A}$ fails to exhibit the required "structural particularities." This is the point I had in mind at footnote [16].

It had become apparent by the time we reached (61) that—the evidence of (54) notwithstanding—a properly drawn abstract of the relationship between Hamiltonian particle mechanics and the Hamiltonian theory of fields reads not

$$H(p, q) \quad \longleftrightarrow \quad \mathcal{H}$$

but

$$H(p, q) \quad \longleftrightarrow \quad H = \int \mathcal{H} dx$$

—as would, in fact, have been evident from the outset had we proceeded by the lattice-refinement technique. But consider: field-theoretic objects of type $H, J_r, A, B, \ldots$ are by nature *functions of functions* [32]—they are, in short, "functionals"—while the notion of a "Poisson bracket" is, in all of its diverse manifestions, rooted in the concept of *differentiation*. The Poisson bracket guards the entry portal to the higher reaches of Hamiltonian mechanics; pretty clearly, if we, as field theorists, are ever to penetrate those higher reaches—are ever, for example, to understand the deeper meaning of (61)—we must be in possession of a "functional calculus." That, therefore, is a topic to which I promise to return. But for the moment it seems to me advisable to take temporary leave of theory-building in order to explore what the results already in hand have to say about illustrative concrete cases.

[32] They are, more precisely, number-valued functions of *sets* of functions $\{\pi, \varphi\}$ *and their "spatial" first partials* $\{\nabla\pi, \nabla\varphi\}$.

Examples of the Hamiltonian method at work. Look first to the single-field system

$$\mathcal{L} = \tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\varphi_x^2 - F(\varphi)$$

first encountered on p. 37. There are two independent variables, therefore two distinct variants ($\mathcal{H}^t$ and $\mathcal{H}^x$) of the Hamiltonian formalism. Working first within the former, we write

$$\mathcal{L} = \tfrac{1}{2}\dot{\varphi}^2 - \tfrac{1}{2}\varphi_x^2 - F(\varphi) \quad \text{with} \quad \dot{\varphi} \equiv \varphi_t$$

and introduce

$$\pi = \partial\mathcal{L}/\partial\dot{\varphi} = \dot{\varphi}$$

The function $\dot{\varphi}(\pi, \varphi_x, \varphi)$ is in this case very simple: $\dot{\varphi} = \pi$. The Hamiltonian density is given therefore by

$$\begin{aligned}\mathcal{H} &= \left[\pi\dot{\varphi} - \mathcal{L}(\varphi, \dot{\varphi}, \varphi_x)\right]_{\dot{\varphi}\to\pi} \\ &= \tfrac{1}{2}\pi^2 + \tfrac{1}{2}\varphi_x^2 + F(\varphi)\end{aligned}$$

The canonical equations of motion (54) read

$$\begin{aligned}\dot{\varphi} &= +\left\{\frac{\partial}{\partial\pi} - \frac{\partial}{\partial x}\frac{\partial}{\partial\pi_x}\right\}\mathcal{H} = \pi \\ \dot{\pi} &= -\left\{\frac{\partial}{\partial\varphi} - \frac{\partial}{\partial x}\frac{\partial}{\partial\varphi_x}\right\}\mathcal{H} = -F'(\varphi) + \varphi_{xx}\end{aligned}$$

from which it is very easy to recover the Lagrangian field equation

$$\varphi_{tt} - \varphi_{xx} + F'(\varphi) = 0$$

The Noetherian analysis of p. 28, applied to the specific system now at hand, shows it to be an implication of the assumed t-independence of $\mathcal{L}$ that

$$\partial_t S^t{}_t + \partial_x S^x{}_t = 0$$

while

$$\partial_t S^t{}_x + \partial_x S^x{}_x = 0$$

follows similarly from the assumed x-independence. Explicit descriptions of the quantities

$$\begin{pmatrix} S^t{}_t & S^t{}_x \\ S^x{}_t & S^x{}_x \end{pmatrix} \equiv \begin{pmatrix} S^0{}_0 & S^0{}_1 \\ S^1{}_0 & S^1{}_1 \end{pmatrix}$$

can be read off from (34), but since in the Hamiltonian formalism $\mathcal{H}^t$ we have interest only in $S^t{}_t$ and $S^t{}_x$ we restrict our explicit attention to those; we find

$$\begin{aligned}S^t{}_t &= \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\dot{\varphi} - \mathcal{L} \\ S^t{}_x &= \frac{\partial\mathcal{L}}{\partial\dot{\varphi}}\varphi_x\end{aligned}$$

from which (using $\dot\varphi = \pi$ to eliminate all reference to $\dot\varphi$) we obtain

$$\begin{aligned}\mathcal{S}_t &= \tfrac{1}{2}\pi^2 + \tfrac{1}{2}\varphi_x^2 + F(\varphi)\\ \mathcal{S}_x &= \pi\varphi_x\end{aligned}$$

Since $\mathcal{S}_t$ *is precisely the Hamiltonian density* $\mathcal{H}$, it is trivially the case that $[\,\mathcal{S}_t, \mathcal{H}\,] = 0$, and therefore trivial also that

$$S_t = \int \mathcal{S}_t dx \quad \text{is conserved:} \quad \dot S_t = \int \underbrace{[\,\mathcal{S}_t, \mathcal{H}\,]}_{0} dx = 0$$

By calculation we find, on the other hand, that

$$\begin{aligned}[\,\mathcal{S}_x, \mathcal{H}\,] &= -\pi_x\pi - \varphi_x\big(F'(\varphi) - \varphi_{xx}\big)\\ &= \tfrac{\partial}{\partial x}\mathcal{W} \quad \text{with} \quad \mathcal{W} \equiv \tfrac{1}{2}\varphi_x^2 - \tfrac{1}{2}\pi^2 - F(\varphi)\end{aligned}$$

does *not* vanish, but *has the structure of a divergence*. Therefore

$$\dot S_x = \int \tfrac{\partial}{\partial x}\mathcal{W}\, dx = \text{boundary terms}$$

from which it follows that

$$S_x = \int \mathcal{S}_x dx \quad \text{is conserved if the boundary terms vanish}$$

Since, as we have seen, S_x admits of physical interpretation as the *total linear momentum* of the system, we are not at all surprised to be reminded that "conservation of linear momentum requires that the system be isolated" (no boundary effects).

We follow those same ideas now down a less well-trodden path. Relativistic considerations (which we are not yet in position to entertain, though they are of no great profundity; we note simply that our $\mathcal{L}$ contains $\mathcal{L} = \varphi_{tt} - \varphi_{xx}$ as a special case, and that this is the system which historically served as the cradle of Special Relativity) inspire an interest in the Lorentz-transform properties of the system $\mathcal{L}$. In 2-dimensional spacetime the Lorentz group is a one-parameter group, the elements of which can be described

$$\begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \longrightarrow \begin{pmatrix} X^0 \\ X^1 \end{pmatrix} = e^{\omega\mathbb{G}} \begin{pmatrix} x^0 \\ x^1 \end{pmatrix} \quad \text{with} \quad \mathbb{G} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

To describe an infinitesimal Lorentz transformation we therefore write

$$\begin{aligned}\delta_\omega x^0 &= \mathcal{X}^0\delta\omega \quad \text{with} \quad \mathcal{X}^0 = x^1\\ \delta_\omega x^1 &= \mathcal{X}^1\delta\omega \quad \text{with} \quad \mathcal{X}^1 = x^0\end{aligned}$$

and, to express the presumption that φ transforms as a *scalar* field, set $\Phi = 0$. Returning now to (29) (and reverting to our recent notational conventions: $t \longleftarrow x^0$, $x \longleftarrow x^1$) we obtain the Noetherian current with components

$$\begin{aligned} K^t &= +\varphi_t(\varphi_t x + \varphi_x t) - (\tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\varphi_x^2 - F)x \\ K^x &= -\varphi_x(\varphi_t x + \varphi_x t) - (\tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\varphi_x^2 - F)t \end{aligned}$$

which after simplifications can be written

$$\begin{aligned} K^t &= +\varphi_t\varphi_x t + (\tfrac{1}{2}\varphi_t^2 + \tfrac{1}{2}\varphi_x^2 + F)x \\ K^x &= -\varphi_t\varphi_x x - (\tfrac{1}{2}\varphi_t^2 + \tfrac{1}{2}\varphi_x^2 - F)t \end{aligned}$$

If one writes out $\partial_t K^t + \partial_x K^x$ and draws upon the Lagrangian field equation one discovers without difficulty that in fact

$$\partial_t K^t + \partial_x K^x = 0$$

But from a Hamiltonian point of view—more precisely, from the viewpoint of the $\mathcal{H}^t$ formalism—the object of interest is the observable density

$$\mathcal{K} \equiv K^t\big|_{\varphi_t \longrightarrow \pi} = \pi\varphi_x t + \mathcal{H}x$$

Noting that $\mathcal{K}$ displays some *explicit* t-dependence, we compute

$$\begin{aligned} [\,\mathcal{K}, \mathcal{H}\,] + \frac{\partial \mathcal{K}}{\partial t} &= \big\{F'x - (\pi t)_x - (\varphi_x x)_x\big\}\big\{\pi\big\} - \big\{F' - \varphi_{xx}\big\}\big\{\varphi_x t + \pi x\big\} + \pi\varphi_x \\ &= t \cdot \tfrac{\partial}{\partial x}\mathcal{W} \quad \text{with} \quad \mathcal{W} \equiv \tfrac{1}{2}\varphi_x^2 - \tfrac{1}{2}\pi^2 - F(\varphi) \quad \text{as before} \end{aligned}$$

Therefore

$$\dot{K} = \int \Big\{[\,\mathcal{K}, \mathcal{H}\,] + \tfrac{\partial}{\partial t}\mathcal{K}\Big\}dx = t \cdot \int \tfrac{\partial}{\partial x}\mathcal{W}dx = \text{boundary terms}$$

from which it follows that

$$K = \int \mathcal{K}dx \quad \text{is conserved if the boundary terms vanish}$$

So distinguished is the ancestry of K—it is rooted in, and a symptom of, the Lorentz covariance of our system—that some urgency attaches to the matter of its physical interpretation. The following observations are intended not to resolve that question, but only to indicate the direction in which the resolution surely lies. We begin by noting that $\int(\mathcal{H}x)dx$ is the first moment of the energy distribution. Equivalently, the first moment of the mass distribution. It becomes on this basis fairly natural to consider that

$$X(t) = \int(\mathcal{H}x)dx \quad \text{locates the "center of mass" of the field}$$

In this notation $K = X(0)$, and we have

$$X(t) = X(0) + Vt$$

with $V = -\int(\pi\varphi_x)dx = -\int(\pi\varphi)_x dx + \int(\pi_x\varphi)dx$, where $\int(\pi\varphi)_x dx$ has in fact the nature of a boundary term, and might therefore be dropped. More to the point: if we write $V = \int \mathcal{V}dx$ then by quick calculation we find

$$[\mathcal{V}, \mathcal{H}] = -\tfrac{\partial}{\partial x}\mathcal{V} - \pi_x F(\varphi)$$

from which it follows that V is itself a constant of the field motion, except for effects attributable to $F(\varphi)$. In the absence of such effects, we can, by these interpretations, consider $K = 0$ to be telling us that—however complicated the motion of the field itself may be, its center of mass drifts with constant velocity. In the mechanics of N-particle systems, Galilean covariance can be exploited to similar effect.[33]

Now we take off our Hamiltonian hat $\mathcal{H}^t$, put on the hat $\mathcal{H}^x$, and look to the physics of the same system as before. We agree to retain the convention that the occurance of an "overdot" $\dot{}$ signifies *differentiation with respect to "the parameter,"* and accept that—since "the parameter" is now not t but x—all allusions to "motion" have acquired suddenly a novel meaning. We write

$$\mathcal{L} = \tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\dot{\varphi}^2 - F(\varphi) \quad \text{with} \quad \dot{\varphi} \equiv \varphi_x$$

and assign new meaning to the conjugate momentum field[34]

$$\pi = \partial\mathcal{L}/\partial\dot{\varphi} = -\dot{\varphi}$$

The Hamiltonian density is given in $\mathcal{H}^x$ therefore by

$$\begin{aligned}\mathcal{H} &= \left[\pi\dot{\varphi} - \mathcal{L}(\varphi, \varphi_t, \dot{\varphi})\right]_{\dot{\varphi}\to-\pi} \\ &= -\tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\pi^2 + F(\varphi)\end{aligned}$$

It is, as it was in $\mathcal{H}^t$, a Legendre transform of $\mathcal{L}$, but a *different* Legendre transform. The canonical equations of motion now read

$$\begin{aligned}\dot{\varphi} &= +\left\{\frac{\partial}{\partial\pi} - \frac{\partial}{\partial t}\frac{\partial}{\partial\pi_t}\right\}\mathcal{H} = -\pi \\ \dot{\pi} &= -\left\{\frac{\partial}{\partial\varphi} - \frac{\partial}{\partial t}\frac{\partial}{\partial\varphi_t}\right\}\mathcal{H} = -F'(\varphi) - \varphi_{tt}\end{aligned}$$

[33] See Classical Mechanics (1983), pp. 170 & 253 and especially—because it is much more "field-theoretic" in spirit—the discussion which appears in Classical Electrodynamics (1980) at pp. 323–329.

[34] In order to keep simple things simple, I am asking my reader simply to *remember* that π and similar constructs have now new meanings; to distinguish π^t from π^x seems excessively pedantic, and lends off-putting clutter to the symbols which denote the partial derivatives of those objects.

from which it is as easy as it was before to recover the Lagrangian field equation

$$\varphi_{tt} - \varphi_{xx} + F'(\varphi) = 0$$

We have observed that—and why—the $\mathcal{H}^t$ formalism exhibits a special interest in the t-row of the stress-energy tensor, and are not surprised to discover that the $\mathcal{H}^x$ formalism is similarly enamoured of the x-row. We have

$$\begin{aligned} S^x{}_t &= \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \varphi_t \\ S^x{}_x &= \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \dot{\varphi} - \mathcal{L} \end{aligned}$$

from which we obtain

$$\begin{aligned} \mathcal{S}_t &= \pi \varphi_t \\ \mathcal{S}_x &= -\tfrac{1}{2}\varphi_t^2 - \tfrac{1}{2}\pi^2 + F(\varphi) \end{aligned}$$

Reminding ourselves that the variational derivatives $\frac{\delta}{\delta \varphi}$ and $\frac{\delta}{\delta \pi}$ have now (see again the most recent version of the canonical equations of motion) a new meaning, and that so also does the Poisson bracket, we compute

$$[\,\mathcal{S}_t, \mathcal{H}\,] = \tfrac{\partial}{\partial t}\mathcal{U} \quad \text{with} \quad \mathcal{U} \equiv -\tfrac{1}{2}\varphi_t^2 + \tfrac{1}{2}\pi^2 - F(\varphi)$$

We introduce $S_t = \int \mathcal{S}_t \, dt$ and from

$$\dot{S}_t = \int [\,\mathcal{S}_t, \mathcal{H}\,]\, dt = \text{temporal boundary terms}$$

conclude that S is "conserved"—a constant of the "motion"—if the "temporal boundary terms" vanish. A simpler argument (one has only to notice that $\mathcal{S}_x = \mathcal{H}$, from which $\dot{S}_x = \int [\,\mathcal{S}_x, \mathcal{H}\,]\, dt = 0$ follows trivially) establishes the unconditonal conservation of $S_x = \int \mathcal{S}_x \, dt$.

The $\mathcal{H}^x$ formalism has led us, with an inevitability born of its familiar internal logic, to the perception of a population of conservation laws of this unfamiliar general type:

> Sit in Portland for an eternity, collecting data sufficient to permit the evaluation of (say) $S_x = \int \mathcal{S}_x \, dt$. The result you ultimately obtain is the same as you would have obtained had you instead elected to sit in Pocatello. Or Paris, or Prague.

Such statements are "bizarre" for precisely the reason, and to precisely the extent, that "motion" within $\mathcal{H}^x$ is bizarre: it is x-parameterized, and therefore runs counter to the mechanical experience which has made unwitting "time-Chauvinists" of us all—relativistically untenable though we recognize such a position to be. In particle mechanics our "time-Chauvinism" is, I think, pretty much forced upon us; we may, if we wish, tinker with the metrization of time ($t \longrightarrow \theta = \theta(t)$), but the subject-matter provides no independent

variable fundamentally distinct from the time parameter, and no plausible alternative to such statements as this: "A particle can easily be at the same point at two times, but cannot easily be at two points at the same time." But now we are doing field theory, which provides an abundance of alternatives to t-parameterization, and which treats systems that are *typically* "at two points at the same time." Perhaps, therefore, we should relax the tenacity with which we cling to some entrenched habits of thought. Why, after all, do we collect conservation theorems? For a variety of reasons, of which some are more narrowly technical than others, but principally because *conservation laws nourish physical intuition*; they permit us to see unity in the contingent details

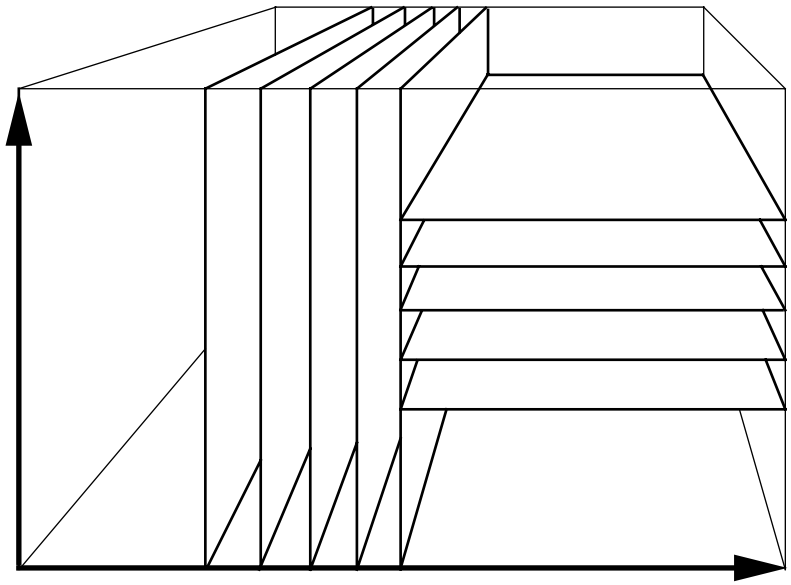


FIGURE 8: *The "standard" Hamiltonian formalism $\mathcal{H}^t$ supplies conservation theorems which are written onto consecutive t-slices, as illustrated on the right half of the figure. The formalism $\mathcal{H}^x$ supplies, on the other hand, theorems which incorporate data written onto consecutive x-slices. It is argued in the text that both contribute usefully to a fuller understanding of the physical system φ, which in itself just "sits there, waiting to be probed" in as many ways as we can think of.*

of particular cases, in the confusing complexity typical of particular solutions of the equations of motion (in field theory: the field equations); they permit us even to say useful things when no solutions are known. It would seem in this light extravagant to discard such "conservation theorems" as $\mathcal{H}^x$ stands ready to provide, simply on the grounds that they come to us nameless, and require of us that we exercise some unfamiliar modes of thought. Nor, in the end, would

such a "policy of abandonment" even be tenable. For if

$$t \longrightarrow T = T(t,x)$$
$$x \longrightarrow X = X(t,x)$$

then the conservation theorems which issue from $\mathcal{H}^T$ will , in general, *conflate* those which issue from $\mathcal{H}^t$ and $\mathcal{H}^x$. As so also, of course, will those which issue from $\mathcal{H}^X$. Directly related to the preceeding remark is this final curious point: We have proceeded from a comparison of the output (so far as relates to conservation theorems) of the $\mathcal{H}^x$ formalism with that of the more familiar $\mathcal{H}^t$ formalism. But there is, in fact, nothing peculiarly "Hamiltonian" about our conclusions, for they were latent already in the Lagrangian formalism, where the conservation law

$$\partial_i J^i = 0$$

is standardly elaborated

$$\partial_0 J^0 + \text{divergence term} = 0$$
$$\text{divergence term} \equiv \quad \bullet \quad + \partial_1 J^1 + \partial_2 J^2 + \cdots + \partial_n J^n$$

but *can*—equally well, if non-standardly—be elaborated

$$\partial_1 J^1 + \text{divergence term} = 0$$
$$\text{divergence term} \equiv \partial_0 J^0 + \quad \bullet \quad + \partial_2 J^2 + \cdots + \partial_n J^n$$

To write the former is to acquire automatic interest in the t-independence of

$$\int J^0 dx^1 dx^2 \cdots dx^n$$

while the latter engenders interest in the x^1-independence of

$$\int J^1 \; dt \; dx^2 \cdots dx^n$$

This simple point—brought naturally and forcibly to our attention by the Hamiltonian formalism—is (almost unaccountably) passed by unnoticed in most surveys of Lagrangian field theory.

Look back again now to the two-field system

$$\mathcal{L} = \tfrac{1}{2}(\alpha_t \varphi - \alpha \varphi_t) - \alpha_x \varphi_x$$

which engaged our interest already on p. 38 for the reason that it yields the diffusion equation $\varphi_t = \varphi_{xx}$ as a field equation.[35] In naive preparation for what we intend to be "work within the $\mathcal{H}^t$ formalism," we write

$$\mathcal{L} = \tfrac{1}{2}(\dot{\alpha}\varphi - \alpha\dot{\varphi}) - \alpha_x \varphi_x$$

[35] Recall that we had no initial interest in the auxiliary field α, which was introduced as a formal crutch, and was found to satisfy the "backwards (or time-reversed) diffusion equation."

and introduce the conjugate momenta

$$\pi \equiv \frac{\partial \mathcal{L}}{\partial \dot{\varphi}} \quad \text{and} \quad \beta \equiv \frac{\partial \mathcal{L}}{\partial \dot{\alpha}}$$

—only to find that we have run smack into a brick wall. For the resulting equations

$$\pi = -\tfrac{1}{2}\alpha$$
$$\beta = +\tfrac{1}{2}\varphi$$

show that the conjugate fields $\{\pi, \beta\}$ are redundant with their companion fields $\{\varphi, \alpha\}$. One cannot by inversion obtain equations of the form

$$\dot{\varphi} = \dot{\varphi}(\text{fields, conjugate fields, non-parametric partials of those})$$
$$\dot{\alpha} = \dot{\alpha}(\text{fields, conjugate fields, non-parametric partials of those})$$

The procedure

$$\mathcal{L} \xrightarrow[\text{Legendre transformation}]{} \mathcal{H}$$

cannot be carried out, for essentially the reason that the Hessian vanishes:

$$\begin{vmatrix} \partial^2\mathcal{L}/\partial\dot{\varphi}\partial\dot{\varphi} & \partial^2\mathcal{L}/\partial\dot{\varphi}\partial\dot{\alpha} \\ \partial^2\mathcal{L}/\partial\dot{\alpha}\partial\dot{\varphi} & \partial^2\mathcal{L}/\partial\dot{\alpha}\partial\dot{\alpha} \end{vmatrix} = 0$$

But consider: if we take

$$\mathcal{H} = -\varphi_x \pi_x$$

to be (within the $\mathcal{H}^t$ formalism) the Hamiltonian density of a *one*-field system, then the canonical equations

$$\dot{\varphi} = +\left\{ \frac{\partial}{\partial \pi} - \frac{\partial}{\partial x}\frac{\partial}{\partial \pi_x} \right\}\mathcal{H} = +\varphi_{xx} \quad : \quad \text{diffusion equation}$$
$$\dot{\pi} = -\left\{ \frac{\partial}{\partial \varphi} - \frac{\partial}{\partial x}\frac{\partial}{\partial \varphi_x} \right\}\mathcal{H} = -\pi_{xx} \quad : \quad \text{backwards diffusion equation}$$

reproduce precisely the equations formerly obtained as field equations from the Lagrangian density $\mathcal{L}$! What we have in hand—since clearly no procedure of the type

$$\mathcal{L} \xleftarrow[\text{Legendre transformation}]{} \mathcal{H}$$

is possible—is a strikingly efficient example of a "free-standing Hamiltonian theory," a theory unsupported by any underlying Lagrangian framework. The absence of such a framework carries with it the implication that we are cut off from (for example) the assistance of Noether's Theorem; when we undertake to construct descriptions of (say) the energy density, or the momentum density—constructions which the equivalent two-field Lagrangian theory supplies straightforwardly—we will be "on our own." But that circumstance we are prepared to accept as an invitation to invention!

There remains yet some juice to be squeezed from the preceeding example, and the last drop is, in a way, the sweetest. By way of preparation, recall that in mechanics of particles it sometimes proves useful—particularly in connection with the theory of canonical transformations[36]—to notice that the canonical equations of motion

$$\dot{q} = +\frac{\partial H}{\partial p}$$
$$\dot{p} = -\frac{\partial H}{\partial q}$$

can be obtained as the "Lagrange equations" from a certain "meta-Lagrangian." For if

$$\boldsymbol{L} \equiv \tfrac{1}{2}(p\dot{q} - \dot{p}q) - H(p,q) = (p\dot{q} - H) + \text{gauge term}$$

then

$$\left\{\frac{d}{dt}\frac{\partial}{\partial \dot{p}} - \frac{\partial}{\partial p}\right\}\boldsymbol{L} = 0 \quad \Longrightarrow \quad \dot{q} = +\frac{\partial H}{\partial p}$$
$$\left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}} - \frac{\partial}{\partial q}\right\}\boldsymbol{L} = 0 \quad \Longrightarrow \quad \dot{p} = -\frac{\partial H}{\partial q}$$

This remark admits straightforwardly of field-theoretic imitation, for if we define

$$\boldsymbol{\mathcal{L}} \equiv \tfrac{1}{2}(\pi\dot{\varphi} - \dot{\pi}\varphi) - \mathcal{H}(\pi,\varphi) \tag{63}$$

we obtain

$$\left.\begin{aligned}\left\{\frac{\partial}{\partial t}\frac{\delta}{\delta \dot{\pi}} - \frac{\delta}{\delta \pi}\right\}\boldsymbol{\mathcal{L}} = 0 \quad &\Longrightarrow \quad \dot{\varphi} = +\frac{\delta \mathcal{H}}{\delta \pi} \\ \left\{\frac{\partial}{\partial t}\frac{\delta}{\delta \dot{\varphi}} - \frac{\delta}{\delta \varphi}\right\}\boldsymbol{\mathcal{L}} = 0 \quad &\Longrightarrow \quad \dot{\pi} = -\frac{\delta \mathcal{H}}{\delta \varphi}\end{aligned}\right\} \tag{64}$$

This is a result of very general importance. But look now again to the particular case $\mathcal{H} = -\pi_x\varphi_x$; we have

$$\begin{aligned}\boldsymbol{\mathcal{L}} &= \tfrac{1}{2}(\pi\dot{\varphi} - \dot{\pi}\varphi) + \pi_x\varphi_x \\ &= -\underbrace{\left[\tfrac{1}{2}(\dot{\pi}\varphi - \pi\dot{\varphi}) - \pi_x\varphi_x\right]}_{\mathcal{L} = \frac{1}{2}(\dot{\alpha}\varphi - \alpha\dot{\varphi}) - \alpha_x\varphi_x \text{ to within notational adjustment}}\end{aligned}$$

We've been "talking prose all our lives, without knowing it!" What we initially imagined to be an unwelcome "auxiliary field" was actually the *conjugate* field, and what we took to be the Lagrangian $\mathcal{L}$ was actually the *meta*-Lagrangian, $\boldsymbol{\mathcal{L}}$. When we lamented the impossibility of proceeding

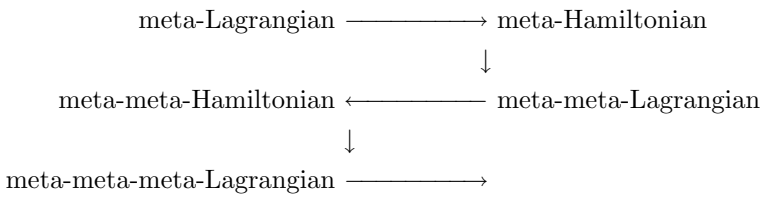
$$\mathcal{L} \xrightarrow[\text{Legendre transformation}]{} \mathcal{H}$$

we were actually commenting on the impossibility of proceeding

$$\boldsymbol{\mathcal{L}} \xrightarrow[\text{Legendre transformation}]{} \boldsymbol{\mathcal{H}}$$

[36] See Classical Mechanics (1983), pp. 217–228.

which is, in fact, cause for rejoicing, for it protects us from the possiblity of infinite regress:



Here I break off this account of the Hamiltonian method, as it applies to the classical theory of fields. My objective has been simply to sketch the lay of the land, and I think that limited objective has been achieved. But clearly, wonders lie just over the horizon.

Classical field theory of a quantum particle. My purpose here will be to illustrate —in what I think readers will agree is a rather surprising context—how the material presently at our command can be used to organize the discussion of some real physics of undeniable importance. In view of that intent, and in order to costume my results in their most familiar dress, I will abandon my formerly somewhat generic mode of expression; I will be careful to insure that all variables are clearly dimensioned, and that all physical parameters and constants are explicitly displayed. Let us agree at the outset that we have equipped ourselves with a "good clock" and an "inertial meter stick;" i.e., that we have coordiniatized spacetime in such a way that the motion of classical free particles can be described $\ddot{x} = 0$. Thus prepared...

To describe the one-dimensional quantum motion of a mass point m in the presence of a potential $U(x)$, we write and undertake to solve—subject to prescribed side conditions and to the normalization condition

$$\int \psi^* \psi dx = 1 \tag{65}$$

—the time-dependent Schrödinger equation

$$-(\hbar^2/2m)\psi_{xx} + U\psi = i\hbar\psi_t \tag{66}$$

The wave function $\psi(x,t)$ is, owing to the presence of the exposed i-factor on the right side of (66), necessarily complex. If, to emphasize that fact, we write

$$\psi = \varphi^1 + i\varphi^2$$

we find that (66) is equivalent to the following coupled *system* of equations

$$\left.\begin{aligned} -(\hbar^2/2m)\varphi^1_{xx} + U\varphi^1 &= -\hbar\varphi^2_t \\ -(\hbar^2/2m)\varphi^2_{xx} + U\varphi^2 &= +\hbar\varphi^1_t \end{aligned}\right\} \tag{67}$$

where evidently it is the "exposed i-factor" that is responsible for the coupling. If, alternatively, we adopt the polar representation

$$\psi = R\exp\left\{\tfrac{i}{\hbar}S\right\}$$

we obtain

$$\left\{-\tfrac{\hbar^2}{2m}\left[R_{xx}+2\tfrac{i}{\hbar}R_xS_x+\tfrac{i}{\hbar}R\left(S_{xx}+\tfrac{i}{\hbar}S_x^2\right)\right]+UR\right\}e^{\frac{i}{\hbar}S}=i\hbar\left(R_t+\tfrac{i}{\hbar}RS_t\right)e^{\frac{i}{\hbar}S}$$

giving—if we abandon the exponential factors, separate the real terms from the imaginary, and make some notational adjustments—

$$\left.\begin{aligned}\tfrac{1}{2m}S_x^2+U+S_t&=\tfrac{\hbar^2}{2m}R^{-1}R_{xx}\\(R^2)_t+(\tfrac{1}{m}S_xR^2)_x&=0\end{aligned}\right\}\tag{68}$$

Or we might, in place of the fields φ^1 and φ^2, elect to work with some linear combination of those fields, writing

$$\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}\varphi^1\\ \varphi^2\end{pmatrix}$$

If, in particular, we elect to set

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}1 & +i\\ 1 & -i\end{pmatrix}$$

we obtain

$$\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}=\begin{pmatrix}\psi\\ \psi^*\end{pmatrix}$$

This little observation lends legitimacy to the frequently very convenient trick whereby one treats ψ and ψ^* as though they were *distinct and independent complex fields*, even though one cannot, in point of fact, change either without inducing change in the other. If we take advantage of this tricky opportunity (as in the future we will consider ourselves free to do casually, without explicit comment), then (66) acquires a co-equal conjugated companion; we have this *uncoupled pair* of Schrödinger equations

$$\left.\begin{aligned}-(\hbar^2/2m)\psi_{xx}+U\psi&=+i\hbar\psi_t\\-(\hbar^2/2m)\psi^*_{xx}+U\psi^*&=-i\hbar\psi^*_t\end{aligned}\right\}\tag{69}$$

It is useful also to note that the Schrödinger equation can, in the case $U=0$, be written

$$\psi_{xx}=D\psi_t\quad\text{with}\quad D\equiv\tfrac{2m}{i\hbar}$$

which has formally the structure of a "diffusion equation (heat equation) with an *imaginary diffusion coefficient*."

The idea now—just crazy enough to be interesting—is to look upon (67), (68) and (69) as providing alternative representations of the *field equations characteristic of a two-component classical (!) field theory*. Semi-arbitrarily electing to work in language provided by the latter representation, we form

$$\left.\begin{aligned} G_1 &\equiv (\hbar^2/2m)\psi_{xx} + i\hbar\psi_t - U\psi \\ G_2 &\equiv (\hbar^2/2m)\psi^*_{xx} - i\hbar\psi^*_t - U\psi^* \end{aligned}\right\} \tag{70}$$

and find it natural to ask: Does there exist a Lagrangian density $\mathfrak{L}$ such that

$$\left.\begin{aligned} G_1 &= \left\{ \frac{\partial}{\partial t}\frac{\partial}{\partial \psi_t} + \frac{\partial}{\partial x}\frac{\partial}{\partial \psi_x} - \frac{\partial}{\partial \psi} \right\}\mathfrak{L} \\ G_2 &= \left\{ \frac{\partial}{\partial t}\frac{\partial}{\partial \psi^*_t} + \frac{\partial}{\partial x}\frac{\partial}{\partial \psi^*_x} - \frac{\partial}{\partial \psi^*} \right\}\mathfrak{L} \end{aligned}\right\} \ ? \tag{71}$$

It was to answer just such questions that the Helmholtz conditions (44) were devised. In the present context

$$N \equiv \text{number of field components} = 2$$
$$m \equiv \text{number of spacetime dimensions} = 2$$

so these are—rather soberingly— $3 + 6 + 1 = 10$ in number. I write them out:

$$\frac{\eth G_1}{\eth \psi^*_{tt}} - \frac{\eth G_2}{\eth \psi_{tt}} = 0 \tag{72.1}$$
$$\frac{\eth G_1}{\eth \psi^*_{tx}} - \frac{\eth G_2}{\eth \psi_{tx}} = 0 \tag{72.2}$$
$$\frac{\eth G_1}{\eth \psi^*_{xx}} - \frac{\eth G_2}{\eth \psi_{xx}} = 0 \tag{72.3}$$

$$\frac{\partial G_1}{\partial \psi_t} + \frac{\partial G_1}{\partial \psi_t} = \frac{\partial}{\partial t}\left[\frac{\eth G_1}{\eth \psi_{tt}} + \frac{\eth G_1}{\eth \psi_{tt}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G_1}{\eth \psi_{xt}} + \frac{\eth G_1}{\eth \psi_{xt}}\right] \tag{73.1}$$
$$\frac{\partial G_1}{\partial \psi^*_t} + \frac{\partial G_2}{\partial \psi_t} = \frac{\partial}{\partial t}\left[\frac{\eth G_1}{\eth \psi^*_{tt}} + \frac{\eth G_2}{\eth \psi_{tt}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G_1}{\eth \psi^*_{xt}} + \frac{\eth G_2}{\eth \psi_{xt}}\right] \tag{73.2}$$
$$\frac{\partial G_2}{\partial \psi^*_t} + \frac{\partial G_2}{\partial \psi^*_t} = \frac{\partial}{\partial t}\left[\frac{\eth G_2}{\eth \psi^*_{tt}} + \frac{\eth G_2}{\eth \psi^*_{tt}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G_2}{\eth \psi^*_{xt}} + \frac{\eth G_2}{\eth \psi^*_{xt}}\right] \tag{73.3}$$
$$\frac{\partial G_1}{\partial \psi_x} + \frac{\partial G_1}{\partial \psi_x} = \frac{\partial}{\partial t}\left[\frac{\eth G_1}{\eth \psi_{tx}} + \frac{\eth G_1}{\eth \psi_{tx}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G_1}{\eth \psi_{xx}} + \frac{\eth G_1}{\eth \psi_{xx}}\right] \tag{73.4}$$
$$\frac{\partial G_1}{\partial \psi^*_x} + \frac{\partial G_2}{\partial \psi_x} = \frac{\partial}{\partial t}\left[\frac{\eth G_1}{\eth \psi^*_{tx}} + \frac{\eth G_2}{\eth \psi_{tx}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G_1}{\eth \psi^*_{xx}} + \frac{\eth G_2}{\eth \psi_{xx}}\right] \tag{73.5}$$
$$\frac{\partial G_2}{\partial \psi^*_x} + \frac{\partial G_2}{\partial \psi^*_x} = \frac{\partial}{\partial t}\left[\frac{\eth G_2}{\eth \psi^*_{tx}} + \frac{\eth G_2}{\eth \psi^*_{tx}}\right] + \frac{\partial}{\partial x}\left[\frac{\eth G_2}{\eth \psi^*_{xx}} + \frac{\eth G_2}{\eth \psi^*_{xx}}\right] \tag{73.6}$$

$$\frac{\partial G_1}{\partial \psi^*} - \frac{\partial G_2}{\partial \psi} = \frac{1}{2}\frac{\partial}{\partial t}\left[\frac{\partial G_1}{\partial \psi^*_t} - \frac{\partial G_2}{\partial \psi_t}\right] + \frac{1}{2}\frac{\partial}{\partial x}\left[\frac{\partial G_1}{\partial \psi^*_x} - \frac{\partial G_2}{\partial \psi_x}\right] \tag{74.1}$$

Taking the G_1 and G_2 as our input, we find by quick calculation that eight of these ten conditions are satisfied—indeed, are trivially satisfied—but that (73.1) and (73.3) are *not* satisfied, for they entail the absurdity $i\hbar = 0$. We are, however, well aware (see again the discussion of this point on p. 34) that the Helmholtz test is (like computer software tends to be) utterly unforgiving of little misalignments of the input data. So enticed are we by the attractiveness of our paradoxical objective, and so encouraged by the knowledge that the closely related forward/backward diffusion equations *can* be made to pass the test, that we experiment a bit... and are led soon enough to *interchange the definitions* (70) of the functions G; we define

$$\left.\begin{aligned} G_2 &\equiv (\hbar^2/2m)\psi_{xx} + i\hbar\psi_t - U\psi \\ G_1 &\equiv (\hbar^2/2m)\psi^*_{xx} - i\hbar\psi^*_t - U\psi^* \end{aligned}\right\} \tag{75}$$

and try again. A few seconds of work is sufficient to establish that the functions G_1 and G_2—thus defined—*do* pass all parts of the Helmholtz test. Thus encouraged to think that it *may* be possible to write (71), we do a little inspired tinkering[37] and are led at length to an $\mathcal{L}$ that works:

$$\mathcal{L} = \tfrac{1}{2}i\hbar(\psi^*_t\psi - \psi^*\psi_t) + \tfrac{\hbar^2}{2m}\psi^*_x\psi_x + \psi^*U\psi \tag{76}$$

The $\mathcal{L}_{\text{Schrödinger}}$ of (76) possesses several striking and important properties to which I would like now to draw attention. It is, in the first place, *manifestly real*

$$\mathcal{L}^* = \mathcal{L}$$

and so also, therefore, is the associated action functional $S_{\mathcal{R}}[\psi] = \int \mathcal{L}\, dxdt$. This is gratifying, for in the contrary case we would have in hand a "multi-component Langrangian" $\mathcal{L} = \mathcal{L}_1 + i\mathcal{L}_2$, and so would have acquired an obligation to venture in to some highly non-standard formal territory.[38] We might also expect (for reasons basic to the theory of functions of several complex variables) to encounter difficulties in giving precise meaning to Hamilton's Principle $\delta S = 0$. Note also the "bilinear" structure of $\mathcal{L}_{\text{Schrödinger}}$; it is (manifestly) linear in the variables $\{\psi, \psi_x, \psi_t\}$ and linear also in the complex conjugates of those variables. Bilinearity accounts for the uncoupled linearity of the Schrödinger equation (69). We observe in passing that if we write $\psi = \varphi^1 + i\varphi^2$ then

[37] In the real world of work-a-day physics one is well-advised to do one's "inspired tinkering" *at the outset*; i.e., to skip the Helmholtzian folderol, which I have written out only for didactic effect. Helmholtzian methods of greatest practical utility when used to establish the *impossibility* of casting the field equations of momentary interest in Lagrangian form; in such applications they alert one to the need to "do something heroic."

[38] See, in this connection, T. Morgan & D. Joseph, "Tensor Lagrangians and generalized conservation laws for free fields," Nuovo Cimento, **39**, 494 (1965) and the literature cited.

$$\mathfrak{L} = \hbar(\varphi^1\varphi^2_t - \varphi^1_t\varphi^2) + \tfrac{\hbar^2}{2m}(\varphi^1_x\varphi^1_x + \varphi^2_x\varphi^2_x) + U(\varphi^1\varphi^1 + \varphi^2\varphi^2) \tag{77}$$

while $\psi = R\exp\left\{\frac{i}{\hbar}S\right\}$ gives

$$\mathfrak{L} = R^2\left[S_t + \tfrac{1}{2m}S_x^2 + U\right] + \tfrac{\hbar^2}{2m}R_x^2 \tag{78}$$

and that the field-variable adjustments which achieve (77) and (78), while they preserve the reality of (76), do violence to its bilinearity. Commentary concerning what is, from some points of view, the most interesting feature of $\mathfrak{L}_{\text{Schrödinger}}$ (namely, that it has actually the status of a *meta*-Lagrangian) will, however, be reserved until the point at issue has been more carefully framed.

Schrödinger's Lagrangian (76) gives rise (by Noetherian analysis) to a stress-energy tensor the components of which can, by (34), be described

$$\begin{aligned}
S^t{}_t &= \frac{\partial\mathfrak{L}}{\partial\psi_t}\,\psi_t + \frac{\partial\mathfrak{L}}{\partial\psi^*_t}\,\psi^*_t - \mathfrak{L} \\
&= -\tfrac{\hbar^2}{2m}\psi^*_x\psi_x - \psi^*U\psi && (79.1)\\
S^x{}_t &= \frac{\partial\mathfrak{L}}{\partial\psi_x}\,\psi_t + \frac{\partial\mathfrak{L}}{\partial\psi^*_x}\,\psi^*_t \\
&= \tfrac{\hbar^2}{2m}(\psi^*_x\psi_t + \psi^*_t\psi_x) && (79.2)\\
S^t{}_x &= \frac{\partial\mathfrak{L}}{\partial\psi_t}\,\psi_x + \frac{\partial\mathfrak{L}}{\partial\psi^*_t}\psi^*_x \\
&= \tfrac{1}{2}i\hbar(\psi^*_x\psi - \psi^*\psi_x) && (79.3)\\
S^x{}_x &= \frac{\partial\mathfrak{L}}{\partial\psi_x}\psi_x + \frac{\partial\mathfrak{L}}{\partial\psi^*_x}\psi^*_x - \mathfrak{L} \\
&= -\tfrac{1}{2}i\hbar(\psi^*_t\psi - \psi^*\psi_t) + \tfrac{\hbar^2}{2m}\psi^*_x\psi_x - \psi^*U\psi && (79.4)
\end{aligned}$$

By straightforward calculation (which draws critically upon the field equations (69); i.e., upon the Schrödinger equation and its conjugate) we find

$$\partial_t S^t{}_t + \partial_x S^x{}_t = -\psi^*\frac{\partial U}{\partial t}\psi \tag{80.1}$$

$$\partial_t S^t{}_x + \partial_x S^x{}_x = -\psi^*\frac{\partial U}{\partial x}\psi \tag{80.2}$$

These equations speak quite intelligibly to us (in language quite consistent with our Noetherian experience): the former says that

> *Energy will be conserved unless the t-dependence of U serves to break the time-translational invariance of the system*,

while the latter says that

> *Momentum will be conserved unless the x-dependence of U serves to break the space-translational invariance of the system.*

Of course, we do not expect momentum to be conserved in the presence of forces $F = -\partial U/\partial x$!

Equations (80) are local statements, very much in the Lagrangian tradition of classical field theory (though we are at present doing the *quantum* mechanics of a *particle*!), and put us in position to state that the total energy of our field system—an "interesting quantity," whether conserved or not—can be described[39]

$$\begin{aligned} E &= \int (-S^t{}_t)\, dx \\ &= \int \left\{ \tfrac{\hbar^2}{2m} \psi^*_x \psi_x + \psi^* U \psi \right\} dx \\ &= \int \left\{ -\tfrac{\hbar^2}{2m} \psi^* \psi_{xx} + \psi^* U \psi \right\} dx \end{aligned}$$

where to achieve the final result I have integrated by parts and discarded the boundary term. Evidently we can, if we wish, write

$$E = \int \psi^* \left\{ \tfrac{1}{2m} \left(\tfrac{\hbar}{i} \tfrac{\partial}{\partial x} \right)^2 + U \right\} \psi \, dx \tag{81}$$

Similarly, the total momentum can be described

$$\begin{aligned} p &= \int S^t{}_x \, dx \\ &= \int \tfrac{1}{2} i\hbar (\psi^*_x \psi - \psi^* \psi_x)\, dx \\ &= -\int i\hbar (\psi^* \psi_x)\, dx \\ &= \int \psi^* \left\{ \left(\tfrac{\hbar}{i} \tfrac{\partial}{\partial x} \right) \right\} \psi \, dx \end{aligned} \tag{82}$$

Readers will not fail to notice that the expressions which stand on the right sides of equations (81) and (82) are precisely the expressions which in quantum mechanics are taken to describe $\langle E \rangle_\psi$ and $\langle p \rangle_\psi$—the *expected values* of energy and momentum, given that the particle is in state ψ. As field theorists we are struck by the fact (which is for us as quantum mechanics already old news!) that (82) can be notated

$$p = \int \psi^* \mathsf{p} \, \psi \, dx \quad \text{where } \mathsf{p} \text{ is the } \textit{differential operator } \mathsf{p} \equiv \tfrac{\hbar}{i} \tfrac{\partial}{\partial x} \tag{83}$$

and that in this notation (81) becomes

$$E = \int \psi^* \mathsf{H} \, \psi \, dx \quad \text{with} \quad \mathsf{H} \equiv \tfrac{1}{2m} \mathsf{p}^2 + U(\mathsf{x}) \tag{84}$$

while (66)—the Schrödinger equation itself—becomes

$$\mathsf{H}\psi = i\hbar \psi_t$$

[39] I exercise here my freedom to introduce a minus sign in order to achieve consistency with notationial and interpretive orthodoxies later on.

The point I have been trying in recent remarks to illustrate is this: if one looks upon the Schrödinger equation with the eyes of a classical field theorist, then one is led *naturally/spontaneously* to the "quantization procedure"

$$\begin{aligned} p &\longrightarrow \mathsf{p} = \tfrac{\hbar}{i}\tfrac{\partial}{\partial x} \\ x &\longrightarrow \mathsf{x} = x\cdot \end{aligned}$$

and to the formulae (such, for example, as (83) and (84)) most characteristic of the quantum theory of particulate motion.

From the (previously noted) bilinear reality of (76) it follows that $\mathfrak{L}$ —whence also $S = \int \mathfrak{L} dx$ —is invariant under "phase transformations" of the form

$$\left.\begin{aligned} \psi &\longrightarrow e^{+i\omega}\psi \\ \psi^* &\longrightarrow e^{-i\omega}\psi^* \end{aligned}\right\} \tag{85.1}$$

and therefore (we note in passing) that (77) is invariant under

$$\left.\begin{aligned} \varphi^1 &\longrightarrow \varphi^1\cos\omega - \varphi^2\sin\omega \\ \varphi^2 &\longrightarrow \varphi^1\sin\omega + \varphi^2\cos\omega \end{aligned}\right\} \tag{85.2}$$

and (78) invariant under

$$\left.\begin{aligned} R &\longrightarrow R \\ S &\longrightarrow S + \hbar\omega \end{aligned}\right\} \tag{85.3}$$

To describe an *infinitesimal* phase transformation we write

$$\left.\begin{aligned} t &\longrightarrow t + \delta_\omega t \quad \text{with} \quad \delta_\omega t = 0\cdot\delta\omega \\ x &\longrightarrow x + \delta_\omega x \quad \text{with} \quad \delta_\omega x = 0\cdot\delta\omega \\ \psi &\longrightarrow \psi + \delta_\omega\psi \quad \text{with} \quad \delta_\omega\psi = +i\psi\cdot\delta\omega \\ \psi^* &\longrightarrow \psi^* + \delta_\omega\psi^* \quad \text{with} \quad \delta_\omega\psi^* = -i\psi^*\cdot\delta\omega \end{aligned}\right\} \tag{86}$$

We have encountered here our first example of a purely "*internal* symmetry" of a multi-component field system; the fields are "folding among themselves," but nothing else is going on. The field variations derive (in the language of p. 24) entirely from "variation of functional form," and not at all from "variation of argument." To the extent that we find (84) an "interesting map" we should, according to Noether, have interest in the current whose components can, by (29), be described

$$\begin{aligned} P^t &= \frac{\partial\mathfrak{L}}{\partial\psi_t}(i\psi) + \frac{\partial\mathfrak{L}}{\partial\psi_t^*}(-i\psi^*) \\ &= \hbar\cdot\psi^*\psi \end{aligned} \tag{87.1}$$

$$\begin{aligned} P^x &= \frac{\partial\mathfrak{L}}{\partial\psi_x}(i\psi) + \frac{\partial\mathfrak{L}}{\partial\psi_x^*}(-i\psi^*) \\ &= \hbar\cdot i\tfrac{\hbar}{2m}(\psi_x^*\psi - \psi^*\psi_x) \end{aligned} \tag{87.2}$$

and we should not be surprised to discover that (in consequence of the field equations)

$$\partial_t P^t + \partial_x P^x = 0$$

The global formulation of this local conservation law reads

$$\int \psi^* \psi \, dx = \text{constant of the field motion} \tag{89}$$

This result acquires interest from several distinct considerations. It shows, in the first place, that no risk of inconsistency was incurred when, at (65), we imposed the normalization condition $\int \psi^* \psi \, dx = 1$. While (89) speaks to our physical intuitions (as historically it spoke in 1925 to Schrödinger himself) of the conservation of some kind of "charge," it does *not* carry within it any hint of anything having to do with a "statistical interpretation of the wave function." It permits—but certainly does not force—us to write (with Born, 1926)

$$\begin{aligned} \text{"probability density"} &= \psi^* \psi \\ &= \varphi^1 \varphi^1 + \varphi^2 \varphi^2 \\ &= R^2 \end{aligned} \tag{90.1}$$

or (collaterally)

$$\begin{aligned} \text{"probability current"} &= i \tfrac{\hbar}{2m} (\psi_x^* \psi - \psi^* \psi_x) \\ &= m \cdot \text{momentum density} \\ &= \tfrac{\hbar}{m} (\varphi_x^2 \varphi^1 - \varphi^2 \varphi_x^1) \\ &= \tfrac{1}{m} S_x R^2 \end{aligned} \tag{90.2}$$

We are in position now to notice, by the way, that the conservation theorem

$$\tfrac{\partial}{\partial t}(\text{probability density}) + \tfrac{\partial}{\partial x}(\text{probability current}) = 0 \tag{91}$$

lends direct physical significance to the 2nd of the "polar Schrödinger equations" (68). It is, finally, of deep interest (and a harbinger of beautiful things to come) that (91) is attributed within the field-theoretic formulation of quantum mechanics to an *internal symmetry* of the theory.

The quantum theory of a particle is (at least as standardly formulated) so profoundly "Hamiltonian" in spirit that we, as field theorists, find it natural to ask: Do the field equations admit of Hamiltonian formulation within $\mathcal{H}^t$? within $\mathcal{H}^x$? Working first within $\mathcal{H}^t$, we look to (76) and obtain

$$\begin{aligned} \text{momentum conjugate to } \psi &= -\tfrac{1}{2} i\hbar \psi^* \\ \text{momentum conjugate to } \psi^* &= +\tfrac{1}{2} i\hbar \psi \end{aligned}$$

Evidently

$$\mathcal{L} \xrightarrow[\text{Legendre transformation}]{} \mathcal{H}$$

is *not possible*, for reasons of precisely the sort discussed already on p. 57. Nor do the variant Lagrangians (77) and (78) afford, in this regard, any advantages. But... suppose we write

$$\pi = \text{momentum conjugate to } \psi = -\tfrac{1}{2}i\hbar\psi^* \tag{92}$$

—which is to make of ψ^* a field "conjugate" to ψ in quite a novel sense— and construe (76) to have the character of a *meta*-Lagrangian, writing (after multiplication by an inconsequential $\frac{1}{2}$)

$$\mathfrak{L} = \tfrac{1}{2}(\pi\psi_t - \pi_t\psi) + \tfrac{i}{\hbar}(\tfrac{\hbar^2}{2m}\pi_x\psi_x + \pi U\psi) \tag{93}$$

Certainly this meta-Lagrangian works, in the easily confirmed sense that

$$\begin{aligned}\frac{\partial}{\partial t}\frac{\partial\mathfrak{L}}{\partial\pi_t} + \frac{\partial}{\partial x}\frac{\partial\mathfrak{L}}{\partial\pi_x} - \frac{\partial\mathfrak{L}}{\partial\pi} &= 0\\ \frac{\partial}{\partial t}\frac{\partial\mathfrak{L}}{\partial\psi_t} + \frac{\partial}{\partial x}\frac{\partial\mathfrak{L}}{\partial\psi_x} - \frac{\partial\mathfrak{L}}{\partial\psi} &= 0\end{aligned}$$

do reproduce the Schrödinger equations (69), which we would now notate

$$\left.\begin{aligned}-\tfrac{\hbar^2}{2m}\psi_{xx} + U\psi &= +i\hbar\psi_t\\ -\tfrac{\hbar^2}{2m}\pi_{xx} + U\pi &= -i\hbar\pi_t\end{aligned}\right\} \tag{94}$$

Comparison of (93) with (63)—which it so closely resembles—leads us to define

$$\mathcal{H}(\pi,\psi) \equiv -\tfrac{i}{\hbar}(\tfrac{\hbar^2}{2m}\pi_x\psi_x + \pi U\psi) \tag{95}$$

and to observe that the implied canonical field equations

$$\left.\begin{aligned}\psi_t &= +\left\{\frac{\partial}{\partial\pi} - \frac{\partial}{\partial x}\frac{\partial}{\partial\pi_x}\right\}\mathcal{H} = -\tfrac{i}{\hbar}(U\psi - \tfrac{\hbar^2}{2m}\psi_{xx})\\ \pi_t &= -\left\{\frac{\partial}{\partial\psi} - \frac{\partial}{\partial x}\frac{\partial}{\partial\psi_x}\right\}\mathcal{H} = +\tfrac{i}{\hbar}(U\pi - \tfrac{\hbar^2}{2m}\pi_{xx})\end{aligned}\right\} \tag{96}$$

again reproduce the Schrödinger equations (94). So quantum mechanics *does*, within $\mathcal{H}^t$, admit of formulation as a Hamiltonian field theory. Note, however, that the Hamiltonian (95) is "free-standing" in the sense that it does not admit of

$$\mathcal{L} \underset{\text{Legendre transformation}}{\longleftarrow\!\longrightarrow} \mathcal{H}$$

What we took at (76) to be a description of $\mathcal{L}_{\text{Schrödinger}}$ was, we now see, actually a deceptively notated description of the meta-Lagrangian $\mathfrak{L}_{\text{Schrödinger}}$.

Within Hamiltonian formalism we attribute the occurance of conservation laws to the special properties acquired by certain Poisson brackets. For example, we might notice that

$$\text{"probability density"} \equiv \psi^*\psi = 2\tfrac{i}{\hbar}\cdot\pi\psi$$

and compute

$$
\begin{aligned}
[\,\pi\psi, \mathcal{H}\,] &= \left\{ \frac{\partial \mathcal{R}}{\partial \psi} - \frac{\partial}{\partial x}\frac{\partial \mathcal{R}}{\partial \psi_x} \right\} \left\{ \frac{\partial \mathcal{H}}{\partial \pi} - \frac{\partial}{\partial x}\frac{\partial \mathcal{H}}{\partial \pi_x} \right\} \\
&\qquad - \left\{ \frac{\partial \mathcal{H}}{\partial \psi} - \frac{\partial}{\partial x}\frac{\partial \mathcal{H}}{\partial \psi_x} \right\} \left\{ \frac{\partial \mathcal{R}}{\partial \pi} - \frac{\partial}{\partial x}\frac{\partial \mathcal{R}}{\partial \pi_x} \right\} \quad \text{with } \mathcal{R} \equiv \pi\psi \\
&= -\tfrac{i}{\hbar} \Big\{ \pi \Big(U\psi - \tfrac{\hbar^2}{2m}\psi_{xx} \Big) - \Big(\pi U - \tfrac{\hbar^2}{2m}\pi_{xx} \Big)\psi \Big\} \\
&= -\tfrac{i}{\hbar}\tfrac{\hbar^2}{2m}(\pi_{xx}\psi - \pi\psi_{xx}) \\
&= -\tfrac{\partial}{\partial x} \Big\{ i\tfrac{\hbar}{2m}(\pi_x\psi - \pi\psi_x) \Big\} \qquad (97)
\end{aligned}
$$

Therefore (discarding a boundary term) $\frac{\partial}{\partial t}\int \pi\psi\, dx = 0$, which is the upshot of (89). We note that the expression interior to the final braces is proportional to "probability current," as described in (90.1). And that if we bring the equations of motion (94) to the braces in the second equation we obtain $[\,\pi\psi, \mathcal{H}\,] = (\pi\psi)_t$; we have in fact obtained by Hamiltonian means an equation which differs from (91) only by an overall factor.

The discovery that the Lagrangian $\mathcal{L}_{\text{Schrödinger}}$ is (at least within $\mathcal{H}^t$) more properly construed to be a meta-Lagrangian carries within it the seed of a very important lesson. For it means that when we successfully (and quite informatively) brought Noetherian methods to bear upon $\mathcal{L}_{\text{Schrödinger}}$ we were actually bringing those methods to bear upon a *meta*-Lagrangian, and that's a much more general object; the arguments of a meta-Lagrangian range not on configuration space but on phase space (which accounts for the fact that such objects are most commonly encountered in connection with the theory of canonical transformations). By (unwittingly) transposing Noether's line of argument from Lagrangian physics to meta-Lagrangian physics we have in effect created a vast *generalization of Noether's method*, as standardly conceived. And since (if I may lapse for a moment into the simpler language of particle mechanics)

$$\mathbf{L}(\dot{p}, \dot{q}, p, q) = \tfrac{1}{2}(p\dot{q} - \dot{p}q) - H(p, q)$$

exists independently of whether or not $H(p, q)$ happens to be "free-standing" (even, that is to say, when an associated $L(\dot{q}, q)$ cannot be constructed), we have in fact resolved the problem which (at the bottom of p. 57) we have described as an "invitation to invention."

We have been witnesses to a conversation between classical field theory and quantum mechanics that seemed a moment ago to be at the point of wrapping up, to the satisfaction of all participants... but which now takes a new turn. For suppose we elect to work (not, as above, within $\mathcal{H}^t$ but) within $\mathcal{H}^x$. We look back again to (76) and obtain

$$\pi \equiv \text{momentum conjugate to } \psi = \tfrac{\hbar^2}{2m}\psi_x^* \qquad (98.1)$$

$$\pi^* \equiv \text{momentum conjugate to } \psi^* = \tfrac{\hbar^2}{2m}\psi_x \qquad (98.2)$$

and observe that it now *is* possible to proceed

$$\mathcal{L} \xrightarrow[\text{Legendre transformation}]{\qquad\qquad\qquad\qquad} \mathcal{H}$$

We write

$$\begin{aligned}\mathcal{H} &= \pi\psi_x + \pi^*\psi_x^* - \mathcal{L} \\ &= \pi\psi_x + \pi^*\psi_x^* - \tfrac{1}{2}i\hbar(\psi_t^*\psi - \psi^*\psi_t) - \tfrac{\hbar^2}{2m}\psi_x^*\psi_x - \psi^*U\psi\end{aligned}$$

and by substitutions $\psi_x \longrightarrow \frac{2m}{\hbar^2}\pi^*$ and $\psi_x^* \longrightarrow \frac{2m}{\hbar^2}\pi$ obtain

$$\begin{aligned}\mathcal{H} &= \mathcal{H}(\pi, \pi^*, \psi, \psi^*) \\ &= \tfrac{2m}{\hbar^2}\pi^*\pi - \psi^*U\psi - \tfrac{1}{2}i\hbar(\psi_t^*\psi - \psi^*\psi_t)\end{aligned} \tag{99}$$

The canonical field equations are now four in number:

$$\psi_x = +\left\{\frac{\partial}{\partial\pi} - \frac{\partial}{\partial t}\frac{\partial}{\partial\pi_t}\right\}\mathcal{H} = \tfrac{2m}{\hbar^2}\pi^* \tag{100.1}$$

$$\pi_x = -\left\{\frac{\partial}{\partial\psi} - \frac{\partial}{\partial t}\frac{\partial}{\partial\psi_t}\right\}\mathcal{H} = U\psi^* + i\hbar\psi_t^* \tag{100.2}$$

$$\psi_x^* = +\left\{\frac{\partial}{\partial\pi^*} - \frac{\partial}{\partial t}\frac{\partial}{\partial\pi_t^*}\right\}\mathcal{H} = \tfrac{2m}{\hbar^2}\pi \tag{100.3}$$

$$\pi_x^* = -\left\{\frac{\partial}{\partial\psi^*} - \frac{\partial}{\partial t}\frac{\partial}{\partial\psi_t^*}\right\}\mathcal{H} = U\psi - i\hbar\psi_t \tag{100.4}$$

The first and third of these equations give back (98). The 1st order canonical equations collectively reproduce precisely the 2nd order Schrödinger equations (94). It should be noticed that when we shift from one variant of Hamiltonian formalism to another we find ourselves writing equations that are superficially identical, but which carry quite distinct meanings, with the result that the associated calculations which can sometimes feel radially different. For example, to account for "probability conservation" (89) within $\mathcal{H}^x$ we compute

$$\begin{aligned}[\psi^*\psi, \mathcal{H}] = &\left\{\frac{\partial\mathcal{R}}{\partial\psi} - \frac{\partial}{\partial t}\frac{\partial\mathcal{R}}{\partial\psi_t}\right\}\left\{\frac{\partial\mathcal{H}}{\partial\pi} - \frac{\partial}{\partial t}\frac{\partial\mathcal{H}}{\partial\pi_t}\right\} \\ &+ \left\{\frac{\partial\mathcal{R}}{\partial\psi^*} - \frac{\partial}{\partial t}\frac{\partial\mathcal{R}}{\partial\psi_t^*}\right\}\left\{\frac{\partial\mathcal{H}}{\partial\pi^*} - \frac{\partial}{\partial t}\frac{\partial\mathcal{H}}{\partial\pi_t^*}\right\} \\ &- \left\{\frac{\partial\mathcal{H}}{\partial\psi} - \frac{\partial}{\partial t}\frac{\partial\mathcal{H}}{\partial\psi_t}\right\}\left\{\frac{\partial\mathcal{R}}{\partial\pi} - \frac{\partial}{\partial t}\frac{\partial\mathcal{R}}{\partial\pi_t}\right\} \\ &- \left\{\frac{\partial\mathcal{H}}{\partial\psi^*} - \frac{\partial}{\partial t}\frac{\partial\mathcal{H}}{\partial\psi_t^*}\right\}\left\{\frac{\partial\mathcal{R}}{\partial\pi^*} - \frac{\partial}{\partial t}\frac{\partial\mathcal{R}}{\partial\pi_t^*}\right\} \quad \text{with } \mathcal{R} \equiv \psi^*\psi \\ = \psi^*\tfrac{\hbar^2}{2m}\pi^* + \psi\tfrac{\hbar^2}{2m}\pi\end{aligned}$$

which by (100.1) and (100.3) can be written

$$= \tfrac{\partial}{\partial x}(\psi^*\psi)$$

This is interesting information, but does not have the form $\frac{\partial}{\partial t}$(something) which

is essential to the establishment of global conservation theorems within the $\mathcal{H}^x$ formalism. But of course! When the Lagrange formalism supplies $\frac{\partial}{\partial t}\rho+\boldsymbol{\nabla}\cdot\boldsymbol{J}=0$ we look within $\mathcal{H}^t$ to $[\rho,\mathcal{H}]$, but within $\mathcal{H}^x$ should look to the x-component of $[\boldsymbol{J},\mathcal{H}]$. Now

$$x\text{-component of ``probability current''} = i\tfrac{\hbar^2}{2m}(\psi^*_x\psi - \psi_x\psi^*)$$

according to (90.2), so drawing upon (98) we have

$$= \tfrac{i}{\hbar}(\pi\psi - \pi^*\psi^*)$$

and by quick computation obtain

$$[\,x\text{-component of ``probability current''},\mathcal{H}\,] = -\tfrac{\partial}{\partial t}(\psi^*\psi) \tag{101}$$

which has precisely the anticipated form. The computational experience which led us to (101) is vividly distinct from that which led us to (97). Similarly distinct—see again FIGURE 8—are the meanings of the conservation theorems thus achieved. For while (97) refers to the invariance of data written onto t-sections in spacetime (and is therefore a "conservation theorem" in the familiar sense), (101) refers to the invariance of (other) data written onto x-sections.

A few comments before I bring this discussion to a close: I begin by drawing attention to the familiar fact that quantum theory is a wonderfully *abstract* physical theory; only by abandoning some of that abstraction do we find ourselves working in the "Schrödinger picture"

$$H|\psi\rangle = i\hbar\tfrac{\partial}{\partial t}|\psi\rangle$$

and only by further arbitrary acts—for example, by electing to work in the x-representation

$$\psi(x) = \langle x|\psi\rangle$$

or perhaps in the associated "Wigner representation"[40]

$$P_\psi(p,x) = \tfrac{2}{h}\int \psi^*(x+\xi)e^{2\frac{i}{\hbar}p\xi}\psi(x-\xi)\,d\xi$$

—does quantum mechanics acquire the status of a field theory. It seems to me unreasonable—unacceptable—that classical field theory should be able to "see," and to comment usefully upon, quantum mechanics only when the latter subject is suitably costumed; if classical field theory has things to say about quantum mechanics, it should be able to say them about "naked" quantum mechanics. The patient-analyst roles have at this point been reversed; quantum mechanics is telling us that to understand her we should first attend to the improvement of classical field theory, at least as it relates to *linear* systems, and (if I read her correctly) that we should attend in particular to the canonical transform

[40] See QUANTUM MECHANICS (1967), Chapter III, pp. 99 *et seq*.

aspects of classical field theory. For it is by transformation theory that we learn to look beyond the representations of things to the structural essentials of things-in-themselves.[41]

Our quantum experience has—relatedly, I think—also alerted us to the richness of the interconnections amongst the Lagrangian, Hamiltonian and meta-Lagrangian modes of conceptualizing the dynamics of specific classical field systems. Recall the pattern of that experience: we started with the $\mathcal{L}$ of (76) and found that

$$\mathcal{L} \longrightarrow \begin{cases} \mathcal{H}^t & \text{is not possible, while} \\ \mathcal{H}^x & \text{is possible} \end{cases}$$

We noticed that a Hamiltonian quantum mechanics of type $\mathcal{H}^t$ did, however, become possible if $\mathcal{L}$ were assigned the meaning of a meta-Lagrangian $\boldsymbol{\mathcal{L}}$; we then obtained this "folliation" of the initial Lagrangian theory:

$$\begin{array}{ccccc} & & \mathcal{H} \text{ within } \mathcal{H}^t & \longleftarrow & \boldsymbol{\mathcal{L}} \\ & \ulcorner & \mathcal{L} & \lrcorner & \\ \mathcal{L} & \longrightarrow & \mathcal{H} \text{ within } \mathcal{H}^x & \longrightarrow & \tilde{\boldsymbol{\mathcal{L}}} \end{array}$$

where

$$\tilde{\boldsymbol{\mathcal{L}}} = \pi\psi_x + \pi^*\psi_x^* - \left\{ \tfrac{2m}{\hbar^2}\pi^*\pi - \psi^* U \psi - \tfrac{1}{2} i\hbar(\psi_t^*\psi - \psi^*\psi_t) \right\}$$

While the phase space associated with the $\mathcal{H}^t$ formalism is 2-dimensional, that associated with the $\mathcal{H}^x$ formalism is 4-dimensional, and might before be expected to support a richer population of symmetry relations; those we might expect to expose (pursuant to a point developed on p. 68) by bringing Noetherian methods to bear on $\tilde{\boldsymbol{\mathcal{L}}}$.

Classical field theory of the Hamilton-Jacobi equation. Schrödinger, as is well known, was led to the equation which bears his name by an avowedly analogical procedure: building upon speculation (1923) of DeBroglie ("the mechanics of particles is in some respects wave-like"), who was himself inspired by Einstein's demonstration (1905) that "the mechanics of electromagnetic radiation is in some respects particle-like," Schrödinger sought (in his phrase) to "complete the optico-mechanical analogy"

$$\begin{array}{ccc} \text{geometrical optics} & \longleftarrow & \text{physical optics} \\ \downarrow & & \downarrow \\ \text{classical mechanics} & \longrightarrow & \boxed{\text{wave mechanics}} \end{array}$$

[41] It was, in fact, precisely thus that quantum mechanics itself *acquired* its abstract identity; see M. Jammer, The Conceptual Development of Quantum Mechanics (1966)

That geometrical optics (the optics of "rays"—a subject implicit in the variational principle of Fermat) and classical mechanics (which addresses the design of the "trajectories" temporally traced by mass points, and is implicit in the variational principle of Lagrange) are homologous structures had been established already by Hamilton, nearly a century before (1830–1834), and rediscovered by Bruns.[42] The idea that geometrical optics is an artifact of an underlying theory of "optical waves"—and becomes available only in a certain approximation[43]—is, of course, so ancient that Newton was at pains to "refute" it, and by the last third of the 19th Century had been brought to a high state of development; Fresnel, Helmholtz, Kirchhoff and others—working in pursuit of an idea original to Huygens[44]—had erected a "physical optics" in which a "wave equation" was the fundamental analytical device, and diffraction, superposition (linearity), interference and polarization were among the most characteristic phenomenological manifestations. In the early 1920's—immediately prior to Schrödinger's involvement—it had become evident that those same phenomena are manifest in the mechanics of small particles. When Schrödinger sought to account for this development by "completion of the optico-mechanical analogy" his more precise objective was to devise a "wave mechanics" which stands to Hamiltonian mechanics (Hamilton-Jacobi theory) as physical optics stands to Hamiltonian optics; it was to the most rarified (and—then as now—least widely

[42] Hamilton himself proceeded from the observation that his formulation of geometrical optics was adaptable to mechanics, where it gave rise to the "canonical equations of motion," the theory of canonical transformations and what we now call "Hamilton-Jacobi theory" (Jacobi entered the picture in 1837). His mechanical work became widely known, but its optical precursor fell rapidly into obscurity—especially in Europe. Almost seventy years were to pass before Heinrich Bruns—a student of Sophus Lie, from whom he had acquired a sensitivity to the *geometry* of differential equations—reported his "discovery" ("Das Eikonal," Sächs. Ber. D. Wiss. **21**,1895) that Hamilton's mechanics could be adapted to the needs of geometrical optics. Contemporary optical theorists generally claim desent from Hamilton, rather than Bruns, but have appropriated Bruns' "eikonal" terminology when referring to what Hamilton called the "characteristic function." Bruns' work, though redundant, did serve to revive interest in Hamiltonian optics, and served to underscore the fact that—on the classical side of the ledger—the "optico-mechanical analogy" refers to an association of the form

$$\text{Hamiltonian optics} \quad \longleftrightarrow \quad \text{Hamiltonian mechanics}$$

[43] What approximation? The essential clue is provided by the fact that while the concept of "ray" is recommended to our intuition by the evidence of our eyes, our acoustic experience does not make it natural to speak of "sound rays."

[44] *Traité de la lumière* (1690). For a suberb account of the developments to which I allude, see B. B. Baker & E. T. Copson, *The Mathematical Theory of Huygens' Principle* (1950).

understood) formulation of classical mechanics that Schrödinger looked for his point of departure.[45] More specifically...

The classical object to which Schrödinger looked for quantum guidance was the Hamilton-Jacobi equation, which in the simplest instance (mass point m moving one-dimensionally in the presence of an impressed potential) can be written

$$\tfrac{1}{2m}S_x^2 + U + S_t = 0 \quad : \quad S \equiv S(x,t) \tag{102.1}$$

and in progressively more general classes of cases becomes

$$\begin{array}{rcll}
\tfrac{1}{2m}\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S + U + S_t = 0 & : & S \equiv S(\boldsymbol{x},t) & : \quad \text{relax one-dimensionality} \\
& \downarrow & & \\
H(\boldsymbol{\nabla}S, \boldsymbol{x}) + S_t = 0 & : & \text{relax restriction on structure of } H(\boldsymbol{p},\boldsymbol{x}) & \\
& \downarrow & & \\
H\big(\tfrac{\partial S}{\partial q}, q\big) + \tfrac{\partial S}{\partial t} = 0 & : & S \equiv S(q,t) \equiv S(\underbrace{q^1, q^2, \ldots, q^n}_{\text{generalized coordinates}}, t) &
\end{array} \tag{102.2}$$

It is a remarkable fact that that the classical mechanics of a particle (or system of particles), when formulated at such a high level[46] of abstraction, is described no longer by a system of coupled ordinary differential equations, but by a single *partial* differential equation; it has, in short become a field theory. Equations of the form

$$S(q,0) = \text{constant, call it } \sigma$$

[45] He broke new ground in the manner in which he did so, but not in the mere fact that he did so; "quantization conditions" of the form $\oint p\,dq = nh$ began to appear at an early point in the history of the "old quantum mechanics," and in 1916 Karl Schwartzschild remarked that such expressions could in every instance be associated with the "action-angle" variables which are the hallmark of the "Hamilton-Jacobi theory of periodic systems" devised by the celestial mechanic C. E. Delauney in 1846. Paul Ehrenfest was at that same time promoting the view that "Bohr-Sommerfeld quantization conditions" should be associated with certain "adiabatic invariants," and was inspired by Schwartzschild's paper to suggest to Jan Burgers, his thesis student, that it might be possible to develop a "Hamilton-Jacobi theory of adiabatic invariants with quantum mechanical applications." Which Burgers promptly did (though Ehrenfest himself claimed never to have understood the work). The perception that Hamilton-Jacobi theory had important things to say about the quantum world became (at least within a small circle of theorists) progressively more widespread, with the curious result that two of the best accounts of classical Hamilton-Jacobi theory (I refer to Max Born's *The Mechanics of the Atom* (1924) and George Birtwhistle's *The Quantum Theory of the Atom* (1925)) were published with quantum intent just in time to be rendered obsolete by Schrödinger. For further discussion and references, see CLASSICAL MECHANICS (1983), pp. 382–423, especially 418–421.

[46] Which is to say: at such a profoundly *geometrical* level...

describe the initial design of a σ-parametrized population of non-intersecting *surfaces* in configuration space. The Hamilton-Jacobi equation describes the $H(q,p)$-generated *motion* of such surfaces.[47] The fundamental relation

$$\boldsymbol{p}(\boldsymbol{x},t) = \boldsymbol{\nabla} S(\boldsymbol{x},t) \tag{103.1}$$

—more generally

$$p_i(q,t) = \tfrac{\partial}{\partial q^i} S(q,t) \quad : \quad i = 1,2,\ldots,n \tag{103.2}$$

—directs our attention to a dual construction: the population of *curves* which (at time t) thread everywhere normally through the population of S-surfaces. It is through that portal and another—the observation that the functions $p_i(q,t)$ associate p-values with q-values, and serve therefore to inscribe (at time t) a *surface on phase space*, a point-set which is subsequently swept along by the $H(q,p)$-generated phase flow—that one recovers contact with the moving-point imagery of (respectively) the Lagrangian and canonical Hamiltonian formalisms; I will soon have occasion to spell out the meaning this remark in somewhat finer detail.

It is, it seems to me, quite natural to ask (and therefore surprising that, so far as I am aware, no one else has previously thought to ask) whether the Hamilton-Jacobi equation—looked upon as a field equation—can be displayed as an instance of a Lagrangian field theory. Here the Helmholtz conditions (44) come to our aid. Looking to the case (102.1), we have a single field ($N = 1$) written on a spacetime of $m = 2$ dimensions, so the Helmholtz conditions are two in number; they read $\partial G/\partial S_t = 0$ and $\partial G/\partial S_x = 0$; i.e.,

$$1 = 0 \quad \text{and} \quad \tfrac{1}{m} S_x = 0$$

which are absurd: *no Lagrange density exists* which would yield (102.1) as a field equation.

This is an unsettling development, for Hamilton-Jacobi theory is a field theory with a uniquely strong claim to "deep and universal significance;" it seems to me unreasonable on its face that such a theory should be denied access to the rich formal resources of Lagrangian field theory. I interpret this development to be a "call to invention;" we must deepen the channel if we are to get this boat afloat. I describe now two distinct strategies for accomplishing precisely that objective:

FIRST STRATEGY

What might be called the "auxiliary variable trick"[48] is always available, but merits serious attention only in cases where *direct physical significance* can be assigned to the auxiliary field or fields, and to the associated field equations.

[47] It is natural to associate such surfaces with "wavefronts," though—for us as for Huygens—such wavefronts possess no undulatory aspect; we find our pre-quantum mechanical selves contemplating "wavefronts without waves."

[48] See again pp. 39–40.

That we have in hand just such a case will emerge, but only after we have introduced new ideas into the Hamilton-Jacobi formalism itself.[49]

Writing

$$\mathcal{L} = A \cdot \left\{ \tfrac{1}{2m} S_x^2 + U + S_t \right\}$$

we observe that, of dimensional necessity,

$$[A] = \frac{1}{(\text{length})^{\text{spatial dimension}}} = \frac{1}{\text{volume}} = \text{density}$$

and obtain field equations

$$\left\{ \partial_t \frac{\partial}{\partial A_t} + \partial_x \frac{\partial}{\partial A_x} - \frac{\partial}{\partial A} \right\} \mathcal{L} = 0 \quad \text{giving} \qquad \tfrac{1}{2m} S_x^2 + U + S_t = 0$$

$$\left\{ \partial_t \frac{\partial}{\partial S_t} + \partial_x \frac{\partial}{\partial S_x} - \frac{\partial}{\partial S} \right\} \mathcal{L} = 0 \quad \text{giving} \quad \partial_t A + \partial_x \left(\tfrac{1}{m} A S_x \right) = 0$$

The latter equation—which in the 3-dimensional case reads

$$\partial_t A + \boldsymbol{\nabla} \cdot \left(\tfrac{1}{m} A \, \boldsymbol{\nabla} S \right) = 0 \tag{104}$$

—has the form of a *continuity equation*, and therein lies the clue to its meaning. It is to make clear that meaning that we have need of the "new ideas" to which I just referred, and it is to acquire those that I now digress:

Let state points be sprinked onto phase space with initial density given by $P(\boldsymbol{x}, \boldsymbol{p}, 0)$. Those points—transported by the ambient phase flow—have by time t achieved the distribution $P(\boldsymbol{x}, \boldsymbol{p}, t)$. Phase flow is symplectic, therefore

[49] The following discussion has been adapted from material which appears on pp. 435–471 and 489–495 in CLASSICAL MECHANICS (1983). A word concerning my motivation in that work: Feynman's sum-over-paths formulation of quantum mechanics assigns central importance—this is a point first recognized by Pauli and others in 1950/51, a point which seems never to have aroused the interest of Feynman himself—to an object called the *Van Vleck determinant*. That object first appears in a paper ("The correspondence principle in the statistical interpretation of quantum mechanics," PNAS **14**, 178 (1928)) concerned with quantum fundamentals, but is itself entirely classical. My objective was to establish the sense in which the Van Vleck determinant

$$D(\boldsymbol{x}, t; \boldsymbol{x}_0, t_0) \equiv \det \left\| \frac{\partial^2 S(\boldsymbol{x}, t; \boldsymbol{x}_0, t_0)}{\partial x^i \partial x_0^j} \right\|$$

—a construction which has already other work to do within the deeper reaches of Hamilton-Jacobi theory (see p. 259 in V. I. Arnold's *Mathematical Methods of Classical Mechanics* (1989))—acquires a "natural predisposition" to assume its quantum mechanical burden.

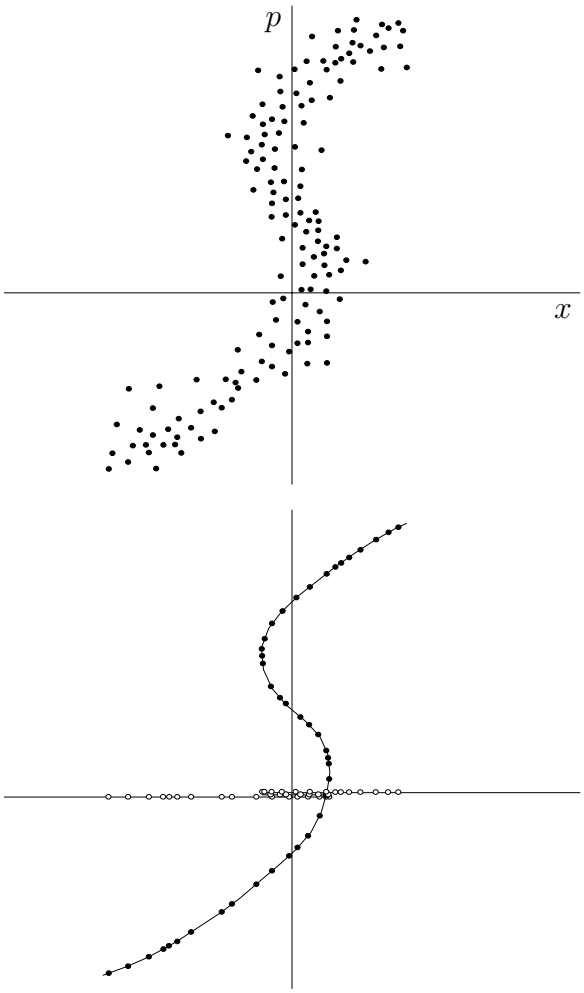


FIGURE 9: *Shown above is an unspecialized distribution $P(x, p, t)$ of state points, such as might be encountered in a statistical mechanical argument. The motion of the distribution is described by (105). In the lower figure the state points have been sprinkled on the surface Σ_t defined by $p(x, t) = S_x(x, t)$. Open circles on the x-axis represent projective images of those points, distributed as described by $A(x, t)$. The distribution in phase space has the specialized structure (106), and $A(x, t)$ satisfies the continuity equation developed in the text. The x-axis is folded to reflect implications of the fact that on a central interval the function $S_x(x, t)$ is triple-valued.*

volume-preserving; therefore the point density in the neighborhood of any co-flowing point is seen to be constant:

$$\begin{aligned}\tfrac{d}{dt}P &= \sum_k \Big\{ \tfrac{\partial P}{\partial x^k}\dot{x}^k + \tfrac{\partial P}{\partial p_k}\dot{p}_k \Big\} + \tfrac{\partial P}{\partial t} \\ &= \sum_k \Big\{ \tfrac{\partial P}{\partial x^k}\tfrac{\partial H}{\partial p_k} - \tfrac{\partial P}{\partial p_k}\tfrac{\partial H}{\partial x^k} \Big\} + \tfrac{\partial P}{\partial t} = 0 \\ &\qquad\qquad\qquad\qquad\qquad\qquad \downarrow \\ &\qquad\qquad\qquad\qquad [\,P, H\,] + \tfrac{\partial P}{\partial t} = 0 \end{aligned} \tag{105}$$

which is the upshot of *Liouville's theorem*.[50] In the "pointwise conception" of Hamiltonian mechanics one uses the canonical equations

$$\dot{x}^k = +\tfrac{\partial}{\partial p_k}H \quad \text{and} \quad \dot{p}_k = -\tfrac{\partial}{\partial x^k}H$$

to watch the motion of individual points in phase space, while in the "global conception" one uses (105) to watch the motion of arbitrarily distributed point *populations*. Hamilton-Jacobi gives rise to an "intermediate conception" if one sprinkles state points on the *surface*—call it Σ_t—which (103) serves to inscribe on phase space; to do so is in effect to assume that the distribution function $P(\boldsymbol{x},\boldsymbol{p},t)$ has the specialized structure

$$P(\boldsymbol{x},\boldsymbol{p},t) = A(\boldsymbol{x},t)\cdot\delta\big(\boldsymbol{p} - \boldsymbol{\nabla}S(\boldsymbol{x},t)\big) \tag{106}$$

Here $A(\boldsymbol{x},t)$ provides a "projective" account of *how* state points are distributed on Σ_t; from

$$\int P(\boldsymbol{x},\boldsymbol{p},t)\,dp_1 dp_2 \ldots dp_n = A(\boldsymbol{x},t)\cdot(\text{multiplicity of } \boldsymbol{\nabla}S(\boldsymbol{x},t) \text{ at } \boldsymbol{x})$$

we see that $A(\boldsymbol{x},t)$ is closely related to the *marginal distribution* in $\boldsymbol{x}$, and becomes precisely the marginal distribution at points where $\boldsymbol{\nabla}S(\boldsymbol{x},t)$ is single-valued. Figure 9 provides illustration of the sense in which distributions of type (106) are "specialized," of how $A(\boldsymbol{x},t)$ acquires projective significance, and of why multiplicity matters. Introducing (106) into (105) we obtain[51]

$$\begin{aligned}[\,P, H\,] + \tfrac{\partial P}{\partial t} &= \big\{A\delta(p-S_x)\big\}_x H_p - \big\{A\delta(p-S_x)\big\}_p H_x + \big\{A\delta(p-S_x)\big\}_t \\ &= \big\{A_x\tfrac{p}{m} + A_t\big\}\delta\,(p-S_x) \\ &\qquad - A\,\delta'(p-S_x)\big\{S_{xx}\tfrac{p}{m} + U_x + S_{xt}\big\}\end{aligned}$$

[50] See H. Goldstein, *Classical Mechanics* (2nd edition 1980), §9-8.

[51] I find it notationally convenient (since the argument is a little bit intricate) to restore now my former presumption that $n = 1$. For that same reason, I restrict my explicit remarks to the case $H(p,x) = \frac{1}{2m}p^2 + U(x)$, and use subscripts to denote partial derivatives.

It has, however, been observe by Dirac[52] that $x\,\delta(x) = 0$; by differentiation we have $x\,\delta'(x) = -\delta(x)$ whence $x\,\delta'(x-a) = a\,\delta'(x-a) - \delta(x-a)$, with the aid of which we obtain

$$\begin{aligned}
&= \big\{\tfrac{1}{m}(S_{xx}A + S_xA_x) + A_t\big\}\delta(p - S_x) \\
&\qquad - A\,\delta'(p - S_x)\big\{\tfrac{1}{m}S_xS_{xx} + U_x + S_{xt}\big\} \\
&= \big\{(\tfrac{1}{m}S_xA)_x + A_t\big\}\delta(p - S_x) \\
&\qquad - A\,\delta'(p - S_x)\tfrac{\partial}{\partial x}\big\{\underbrace{\tfrac{1}{2m}S_x^2 + U + S_t}_{0 \quad \text{by the Hamilton-Jacobi equation}}\big\} \\
&= 0 \quad \text{by (105)}
\end{aligned}$$

The implication is that if $P(\boldsymbol{x}, \boldsymbol{p}, t)$ is of the form (106) then Liouville's theorem (105) entails that $A(\boldsymbol{x}, t)$ satisfies the continuity equation

$$A_t + (\tfrac{1}{m}S_xA)_x = 0$$

In several-dimensional cases (by an elaboration of the same argument) we recover (104). The fact that $A(\boldsymbol{x}, t)$ invites interpretation as a spatial point density conforms well, by the way, to an earlier remark relating to its enforced physical dimensionality.

We are brought thus to the conclusion that while the Hamilton-Jacobi theory of the textbooks—a *single-field* theory—resists display as a Lagrangian field theory, the associated theory of "populated" Σ_t-surfaces—a *two-field theory*, with (in the simplest instance) field equations

$$\left.\begin{aligned}
\tfrac{1}{2m}\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S + U + \tfrac{\partial}{\partial t}S = 0 \\
\tfrac{\partial}{\partial t}A + \boldsymbol{\nabla}\cdot(\tfrac{1}{m}A\boldsymbol{\nabla}S) = 0
\end{aligned}\right\} \tag{107}$$

—does admit of such display, and derives in fact from a very simple Lagrangian:

$$\mathcal{L} = A \cdot \big\{\tfrac{1}{2m}\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S + U + \tfrac{\partial}{\partial t}S\big\} \tag{108}$$

SECOND STRATEGY

In 1979, when I first considered the problem of bringing Hamilton-Jacobi theory within the rubric of of Lagrangian field theory,[53] I was led—with the "integrating factor trick" (see again p. 34) and the "auxiliary field trick"[48] jointly in mind, and "after a bit of exploratory tinkering"—to construct

$$\mathcal{L}_a = e^{S/a}\Big\{\tfrac{a}{2m}S_xB_x + aB_t - UB\Big\} \tag{109}$$

[52] *Principles of Quantum Mechanics* (4th edition, 1958), (7) in §15.

[53] See "A sense in which classical mechanics 'quantizes itself':" notes for a seminar presented 21 January 1980 at Portland State University—appended to CLASSICAL FIELD THEORY (1979).

where a is a constant of (for the moment) arbitrary value but the *dimensionality of action* (made necessary because the exponentiated S carries that dimension) and the auxiliary field B is co-dimensional with the previously-encountered auxiliary field A. The Lagrangian (109) yields field equations

$$\Big\{\partial_t \frac{\partial}{\partial B_t} + \partial_x \frac{\partial}{\partial B_x} - \frac{\partial}{\partial B}\Big\}\mathcal{L}_a = e^{S/a}\Big\{\tfrac{1}{2m}S_x^2 + U + \tfrac{a}{2m}S_{xx} + S_t\Big\} = 0$$
$$\Big\{\partial_t \frac{\partial}{\partial S_t} + \partial_x \frac{\partial}{\partial S_x} - \frac{\partial}{\partial S}\Big\}\mathcal{L}_a = e^{S/a}\Big\{\tfrac{a}{2m}B_{xx} - B_t + \tfrac{1}{a}UB\Big\} = 0$$

which entail

$$\tfrac{1}{2m}S_x^2 + U + \tfrac{a}{2m}S_{xx} + S_t = 0 \tag{110.1}$$
$$\tfrac{a^2}{2m}B_{xx} + UB - aB_t = 0 \tag{110.2}$$

The latter equation, though formally interesting on account of its linearity, appears at present to have the status of a mere artifact, and—particularly since uncoupled to its companion—to have no direct claim to our attention. The former equation contains an "unwanted S_{xx}-term," but that term disappears in the limit $a \downarrow 0$. Evidently the Hamilton-Jacobi equation can be harvested as fruit of the somewhat odd procedure

$$\lim_{a\downarrow 0}\left\{\partial_t \frac{\partial}{\partial B_t} + \partial_x \frac{\partial}{\partial B_x} - \frac{\partial}{\partial B}\right\}\mathcal{L}_a = 0 \tag{111}$$

Formulation of the higher-dimensional analog of this result poses no problem.

The *deus ex machina* is never a welcome participant in scientific discourse. Good theories, like good machines, accomplish their work with the least number of moving parts. We have learned to hold in contempt theories which rely upon the assistance of crutches, which contain "unused parts," and our contempt is the higher the more "fundamental" the theory purports to be. We are motivated by this train of thought to inquire more closely into possibly deeper significance of the auxiliary field B. Such a program entails that we *suspend the limit procedure* $a \downarrow 0$ and *take seriously the two-field theory implicit in* $\mathcal{L}_a$; it entails that we divert our attention from the (solitary) Hamilton-Jacobi equation itself and look with focused attention to the (uncoupled) *pair* of field equations (111).

How to proceed? The question vitually answers itself the minute one notices that adjustment of what one understands to be the "field function"

$$S \longrightarrow \Psi \equiv e^{S/a}$$

entails

$$a^2\Psi_{xx} = e^{S/a}\big\{S_x^2 + aS_{xx}\big\}$$
$$a\Psi_t = e^{S/a}S_t$$

and permits the Lagrange equation $e^{S/a}$(110.1) to be notated

$$\tfrac{a^2}{2m}\Psi_{xx} + U\Psi + a\Psi_t = 0 \tag{112}$$

Remarkably, we have managed (unwittingly) to *linearize* (110.1), and have achieved an equation which—except for the final sign—mimics the structure of (110.2). That mimicry can be made precise by a very simple device: assume a to be *imaginary*, writing

$$a = -i\hbar$$

where $\hbar$ is for the moment not to be confused with Planck's constant: it is a real variable of arbitrary value, and dimensionality $[\hbar]$ = action. Then (112) reads

$$-\tfrac{\hbar^2}{2m}\Psi_{xx} + U\Psi - i\hbar\Psi_t = 0 \tag{113.1}$$

which by conjugation—note that

$$\Psi = e^{\frac{i}{\hbar}S} \quad \text{has become a } \textit{complex-valued} \text{ field function}$$

—assumes precisely the form

$$-\tfrac{\hbar^2}{2m}\Psi^*_{xx} + U\Psi^* + i\hbar\Psi^*_t = 0$$

of (110.2), which now reads[54]

$$-\tfrac{\hbar^2}{2m}B_{xx} + UB \ \ + i\hbar B_t = 0 \tag{113.2}$$

and can be considered to have arisen (together with (113.1)) from this notational variant of (109):

$$\mathcal{L}_\hbar = -\tfrac{\hbar^2}{2m}\Psi_x B_x - i\hbar\Psi B_t - \Psi UB \tag{114}$$

Beyond this point there are several ways to proceed: it is tempting to *identify* (113.2) with the conjugate of (113.1), but dimensional circumstances stand in the way:

$$\Psi\text{—whence also } \Psi^*\text{—are dimensionless, while } [B] = (\text{volume})^{-1}$$

We might, at (114), circumvent this problem by making formal replacements

$$\left.\begin{array}{l}\Psi \hookrightarrow \psi \\ B \hookrightarrow \psi^*\end{array}\right\} \quad \text{with} \quad [\psi] = [\psi^*] = (\text{volume})^{-\frac{1}{2}}$$

We then obtain (after a physically insignificant overall sign reversal)

$$\begin{aligned}\mathcal{L} &= \tfrac{\hbar^2}{2m}\psi^*_x\psi_x + i\hbar\,\psi^*_t\psi + \psi^*U\psi \\ &\quad\downarrow \\ &= \tfrac{\hbar^2}{2m}\psi^*_x\psi_x + \tfrac{1}{2}i\hbar\,(\psi^*_t\psi - \psi^*\psi_t) + \psi^*U\psi\end{aligned} \tag{115}$$

where the point of the final manipulation has been (by gauge transformation) to restore the manifest *reality* of $\mathcal{L}$.

[54] Note that the following equation forces the complexification also of the B-field.

Continuing in the rhetorical pretense that we are "classical physicists who don't already know quantum mechanics," we might argue that the theory latent in (115) is "so pretty that it might be right." But what to do about $\hbar$? Most systems do not supply constants/parameters sufficient to permit assembly of a "natural action." To lend the theory a universality commensurate with its beauty we would be forced to assign $\hbar$ the status of a "constant of Nature" ... and ultimately the observational value

$$\hbar = 1.054592 \times 10^{-27}\text{erg-seconds}$$

We have at (115) recovered precisely the Lagrangian which in discussion subsequent to (76) was found to give rise (with the assistance of Noether's theorem) to all formal aspects of quantum-mechanics-according-to-Schrödinger, though not to enforce the *interpretive* aspects of the latter theory.[55] The classical mechanics of a particle—when written as a (Hamilton-Jacobi) field theory and forced (by these means) into conformity with the general rubric of Lagrangian field theory—has, in effect, "quantized itself." At what point did classical mechanics become something other than it was? Several distinct actions—each seemingly slight, and formally natural in itself—contributed to this development:

- We left classical mechanics behind when we agreed to suspend the limiting process $a \downarrow 0$;
- We went further afield when at $a \to -i\hbar$ we complexified the parameter a, which entailed complexification also of $\Psi = e^{S/a}$ and of B; a 2-field theory had at that point become a 4-field theory.

By that point we had achieved a theory which was *neither* classical mechanics nor quantum mechanics; to achieve the latter (a 2-field theory) we had finally to make the replacements $\Psi \hookrightarrow \psi, B \hookrightarrow \psi^*$.

I now reverse the trend of the argument; taking (115) as my point of departure, I use field-theoretic methods to proceed back again toward the Hamilton-Jacobi theory from which we came:

$$\text{Hamilton-Jacobi theory} \longleftarrow \text{quantum mechanics}$$

We make essential use of the polar representation

$$\psi = \mathsf{R}\exp\left\{\tfrac{i}{\hbar}\mathsf{S}\right\} \tag{116}$$

and will find that most of the work has already been done; the exercise will, however, serve to establish a bridge of the form

[55] It is from those "interpretive aspects"—due to Born—that we acquire rationale for the imposition of *boundary conditions* upon ψ, and also the theory of measurement which becomes the source of (amongst other things) our *initial value* assertions.

FIRST STRATEGY $\longleftrightarrow$ SECOND STRATEGY

and to illuminate the theoretical placement of the so-called "de Broglie-Bohm interpretation of quantum mechanics."

In the discussion culminating in (91) we found it to be (by Noether's theorem) an implication of the *phase-invariance* of $\mathcal{L}$ that

$$\partial_t P + \boldsymbol{\nabla}\cdot\boldsymbol{P} = 0 \quad : \quad \left\{ \begin{array}{l} \text{standardly interpreted to} \\ \text{express "probability conservation"} \end{array} \right.$$

where in polar notation (see again (90)) $P = \mathsf{R}^2$ and $\boldsymbol{P} \equiv \frac{1}{m}\mathsf{R}^2\boldsymbol{\nabla}\mathsf{S}$. So we have

$$\downarrow$$

$$\partial_t(\mathsf{R}^2) + \boldsymbol{\nabla}\cdot(\tfrac{1}{m}\mathsf{R}^2\boldsymbol{\nabla}\mathsf{S}) = 0 \tag{117.1}$$

which bears a striking resemblance to (104), to which it would revert upon substitutional transformation

$$\mathsf{R}^2 \hookrightarrow A \quad \text{and} \quad \mathsf{S} \hookrightarrow S$$

We are, in the light of this observation, not surprised to recall from (78) that introduction of (116) into the higher-dimensional generalization of (115) yields a Lagrangian

$$\mathcal{L} = \mathsf{R}^2 \cdot \big[\mathsf{S}_t + \tfrac{1}{2m}\boldsymbol{\nabla}\mathsf{S}\cdot\boldsymbol{\nabla}\mathsf{S} + U\big] + \tfrac{\hbar^2}{2m}\mathsf{R}_x^2$$

which has *except for the "dangling term"* precisely the structure of (108), and gives rise to field equations (117.1) and

$$2\mathsf{R}\big[\tfrac{1}{2m}\boldsymbol{\nabla}\mathsf{S}\cdot\boldsymbol{\nabla}\mathsf{S} + U + \mathsf{S}_t\big] = \tfrac{\hbar^2}{m}\nabla^2\mathsf{R}$$

$$\downarrow$$

$$\tfrac{1}{2m}\boldsymbol{\nabla}\mathsf{S}\cdot\boldsymbol{\nabla}\mathsf{S} + \big\{U - \tfrac{\hbar^2}{2m}\tfrac{1}{\mathsf{R}}\nabla^2\mathsf{R}\big\} + \tfrac{\partial}{\partial t}\mathsf{S} = 0 \tag{117.2}$$

which—except for a solitary extra term of order $O(\hbar^2)$—possess precisely the structure of the Hamilton-Jacobi system (107).

It will be appreciated that equations (117)—their classical appearance notwithstanding—refer to the *quantum* dynamics of a particle; taken together, they are equivalent to (simply a notational variant of) the Schrödinger equations

$$-\tfrac{\hbar^2}{2m}\nabla^2\psi + U\psi = i\hbar\partial_t\psi \quad \text{and its complex conjugate} \tag{118}$$

The classical dynamics of a particle (Hamilton-Jacobi formulation: (107)) is recovered if, in the spirit of several-dimensional WKB theory,[56] one introduces

$$\begin{aligned} \mathsf{S} &= \ S \ + \hbar\mathsf{S}_1 + \hbar^2\mathsf{S}_2 + \cdots \\ \mathsf{R} &= \sqrt{A} + \hbar\mathsf{R}_1 + \hbar^2\mathsf{R}_2 + \cdots \end{aligned}$$

into (117) and retains only the terms of 0^{th} order.

[56] See QUANTUM MECHANICS (1967), Chapter I, p. 88–106.

Looking to the "field-theoretic reconstruction of quantum mechanics," as outlined above, we see that the theory acquired its undulatory aspects—became literally a "wave mechanics"—from the i introduced at $a \to -i\hbar$.[57] Only with those undulatory aspects in place does it become possible (with de Broglie) to entertain associations of the form

$$\text{wavelength} \sim \hbar/\text{momentum}$$
$$\text{frequency} \sim \text{energy}/\hbar$$

Surfaces of constant S admit literally of interpretation as "wavefronts: surfaces of constant phase;" in the limit $\hbar \downarrow 0$ they become surfaces of constant S, the spooky "wavefronts without waves" which Hamilton-Jacobi theory leads one to contemplate.[47]

The connection between "the field theory we call quantum mechanics" and "the field theory we call Hamilton-Jacobi theory" is so richly deep that it is encountered at every turn, and in many guises, whenever one looks either to "the quantization problem" or to its obverse, "the classical limit (correspondence) problem." Schrödinger himself worked from one perception of that connection, Feynman from another. A particularly beautiful account of the intrusion of Hamilton-Jacobi theory into quantum mechanics was devised by Whittaker in some early work[58] which attracted little attention at the time and is today almost forgotten, though in several respects it was anticipatory both of Feynman and of Schwinger. Diverse beasts drink at the same waterhole. On this occasion I will digress to discuss only one of those:

In 1952 David Bohm published the first[59] of series of papers in which he took the Schrödinger equation (117≡118) as his point of departure, but imported from classical mechanics an equation

$$\boldsymbol{p} = \boldsymbol{\nabla}\mathsf{S} \tag{119}$$

imitative of (103). By the latter strategem he made available to quantum mechanics a physical image—populated surface moving in phase space—basic to classical Hamilton-Jacobi theory (see again Figure 9). Given a solution ψ of (118)—and therefore already in possession of all the testable statements his theory is capable of providing—Bohm proceeds to "play classical mechanics"

---

[57] This i is responsible also for the fact that we had to make a "physically insignificant. . . sign reversal" in order to achieve (115). Such a reversal sends $p = \partial\mathcal{L}/\partial q \to -p$, and will be "physically insignificant" only if the physics in question is *time-reversal invariant*.

[58] E.T.Whittaker, "On Hamilton's principal function in quantum mechanics," Proc. Roy. Soc. Edinburgh **61A**, 1 (1941). For an exposition of Whittaker's work see "The quantum mechanical Hamilton-Jacobi equation" in QUANTUM MECHANICS (1967), Chapter 3, pp. 68–83.

[59] "A suggested interpretation of the quantum theory in terms of 'hidden variables,' Parts I & II," Phys. Rev. **85**, 166 & 180. These papers are reprinted in J.A.Wheeler & W.H.Zurek, *Quantum Theory and Measurement* (1983).

in order to gain what he imagines to be a philosophical advantage. By the phrase "play classical mechanics" I mean this: he elects to construe (117) as a specialized mutant of (107). "Specialized" in this regard: to the "point density" function A he assigns the ψ-dictated structure

$$A \xrightarrow[\text{Bohm's specialization}]{} \psi^*\psi = \mathsf{R}^2$$

And "mutant" in this:

$$U \xrightarrow[\text{Bohm's adjustment}]{} U+Q$$
$$Q \equiv -\tfrac{\hbar^2}{2m}\tfrac{1}{\mathsf{R}}\nabla^2\mathsf{R} = -\tfrac{\hbar^2}{2m}\tfrac{1}{2A}\big[\nabla^2 A - \tfrac{1}{2A}\boldsymbol{\nabla}A{\cdot}\boldsymbol{\nabla}A\big]$$
$$\equiv \text{so-called "quantum potential"}$$

In Bohm's quantum-adjusted version of classical mechanics the populated surface Σ_t inscribed on phase space by (119) does *not*—owing to the presence of the "quantum potential"—move classically (except in the limit $\hbar \downarrow 0$), and in one respect its motion is qualitatively quite *un*classical: the motion of Σ_t *depends upon the density A with which state points have been sprinkled upon it*.[60] That detail notwithstanding, Bohm promotes the view that the particle trajectories which emerge from the system

$$\left.\begin{aligned} -\tfrac{\hbar^2}{2m}\nabla^2\psi + U\psi &= i\hbar\partial_t\psi \\ \tfrac{1}{2m}\boldsymbol{\nabla}\mathsf{S}{\cdot}\boldsymbol{\nabla}\mathsf{S} + \big\{U+Q\big\} + \tfrac{\partial}{\partial t}\mathsf{S} &= 0 \\ Q \text{ constructed from } \psi, \text{ as explained above} & \\ \boldsymbol{p} = \boldsymbol{\nabla}\mathsf{S} & \end{aligned}\right\} \qquad (120)$$

are as "objectively real" as those which emerge from the Hamilton-Jacobi system (107), and that ψ provides representation of a force field as objectively real as (say) the electromagnetic field. A detailed account of the resulting "causal interpretation of quantum mechanics" (the outlines of which were anticipated already by de Broglie in 1932) can be found in a recent monograph by Peter Holland,[61] and a broad selection of more narrowly focused essays—some by major figures (Roger Penrose, Berhard d'Espagnat, Richard Feynman, Y. Aharonov, J. S. Bell and others), some by assorted philosophers and psychiatrists—has been edited by B. J. Hiley & F. David Peat.[62]

---

[60] Bohm himself might have taken exception to my image of "state points sprinkled on Σ_t. He imagined himself to be describing the motion of a *single* state point (or particle), and the intrusion of A into the theory to be "merely a consequence of our [unavoidable and irreducible] ignorance or the precise initial conditions of the particle." In my view the distinction is mainly one of language and emphasis, though it is for Bohm the source of his allusion to "hidden variables."

[61] *The Quantum Theory of Motion* (1993). Holland's §2.6—entitled "Classical mechanics as a field theory" (see also his equation (8.14.19))—bears a striking similarity to material presented in recent pages, though his point of departure and intended destination are both quite different from my own.

[62] *Quantum Implications: Essays in Horour of David Bohm* (1987).

The question "Is quantum mechanics complete?" was debated to a virtual standstill by Bohr and Einstein during the years 1925–31,[63] and brought (or so it was imagined) to a kind of conclusion (in the affirmative) by von Neumann, whose "impossibility proof"[64] appeared in 1932. That the question is today more alive than ever, and exploration of the "foundations of quantum mechanics" has become a cottage industry, is due in substantial part to Bohm. Not that Bohm attracted (except from fringe elements) many converts to his views: most physicists were no more inclined than I am to optimism that our understanding of quantum mechanics can be deepened in any significant or useful way by the addition of elements which can be defended/rebutted only by philosophical debate, and many found technical fault with Bohm's proposal.[65] Bohm's 1952 papers did, however, draw attention to the fact that von Neumann's assumptions were susceptible to criticism (and, latently, to circumvention), and did lend encouragement to those who sought—for whatever reason—to escape the dogma of the "Copenhagen interpretation;" they did, in short, inspire thought. Thus J. S. Bell, who at one point wrote[66] that "Bohm's... papers were for me a revelation. ... I have always felt that people who have not grasped the ideas of those papers... are handicapped in any discussion of the meaning of quantum mechanics." Y. Aharonov remarks that Bohm's theory is "often accused of artificiality and inelegance, and doubtless it is guilty of both," but continues: "But to make such accusations, and to leave it at that, is to entirely miss the point. What Bohm was after... was not elegance and not naturalness; Bohm's intentions were simply to produce a theory which, *whatever* its other characteristics, had *logically clear foundations*..."[67] Quoting again from Bell: "It is easy to find good reasons for disliking the de Broglie–Bohm picture. Neither de Broglie nor Bohm liked it very much; for both of them it was only a point of departure. Einstein also did not like it very much.

---

[63] See P. A. Schilpp, *Albert Einstein: Philosopher-Scientist* (1951), Chapter 7; Abraham Pais, *'Subtle is the Lord...': The Science and the Life of Albert Einstein*, Chapter 25 or Chapter 5 of Jammer (cited below). It is interesting to recall in this connection that the title of the famous EPR paper, which appeared somewhat later (A. Einstein, B. Podolsky & N. Rosen, Phys. Rev. **47**, 777 (1935)), is "Can the quantum mechanical description of Nature be considered complete?"

[64] J. von Neumann, *Mathematische Grundlagen der Quantenmechanik* (1932), Chapter 4, §§1 & 2.

[65] For a wonderfully detailed account of the issues surrounding Bohm's work and its reception, see Max Jammer, *The Philosophy of Quantum Mechanics: The Interpretations of Quantum Mechanics in Historical Perspective* (1974), Chapter 7, especially §7.5.

[66] "Beables for quantum field theory," *Speakable and unspeakable in quantum mechanics* (1987), reprinted in Hiley & Peat.

[67] The quotation is taken from "The issue of retrodiction in Bohm's theory," which appears in Hiley & Peat. The "Aharanov-Bohm effect" was first described in "Significance of electromagnetic potentials in quantum theory," Phys. Rev. **115**, 485 (1959).

He found it 'too cheap,' although, as Born remarked, 'it was quite in line with his own ideas.' "[68]

I feel I owe my reader an explanation: Why—in a work ostensibly concerned with the classical theory of fields—have I allowed myself the indulgence of the preceding digression? The answer resides in the collision of two circumstances: We found it to be the case that the classical mechanics of a particle, when approached field theoretically from just the right angle, very nearly "quantizes itself," but that the line of argument in question only hints at—and certainly does not enforce—the *interpretive* statements standard to quantum theory. Oz Bonfim—my colleague, and the occupant of the office next to mine—has in recent years been exploring this question: Can the "orbital concreteness" which the Bohm theory introduces into quantum theory be exploited to lend sharper meaning to the notion of "quantum chaos"? It became therefore natural to ask: Does the field theoretic quantization procedure lead to a formalism as "naturally predisposed" to Bohm's "causal interpretation" as to the standard (Bohr/Born) interpretation? Had the answer been in the affirmative, I would have looked upon Bohm's creation with more enthusiasm (which is to say: with less dubiousness) than has been my habit. But I have come to the conclusion that appropriation of the equation

$$\boldsymbol{p} = \boldsymbol{\nabla}\mathsf{S}$$

—an act which lies at the heart of Bohm's program—is field theoretically unmotivated; the equation is a gratuitous import which, since it stands with one leg planted in a formalism dominated by the symplectic group and the other in a formalism dominated by the unitary group, leads to a fantasy at risk of becoming transformation-theoretically unstuck unless managed carefully.

I conclude with Bohm-inspired remarks which culminate in formulation of an open question which seems to me to be of independent field theoretic interest. From (116) it follows (recall (90)) that

$$\begin{aligned}
\mathsf{R}^2 &= \psi^*\psi \\
&= \text{probability density} \\
\boldsymbol{\nabla}\mathsf{S} &= \hbar\boldsymbol{\nabla}\arctan\Big\{i\,\frac{\psi^* - \psi}{\psi^* + \psi}\Big\} \\
&= \tfrac{1}{2}i\hbar\,\frac{\psi\boldsymbol{\nabla}\psi^* - \psi^*\boldsymbol{\nabla}\psi}{\psi^*\psi} \\
&= m\cdot\frac{\text{probability current}}{\text{probability density}}
\end{aligned}$$

This information could have been extracted directly from (117.1), but I find the present argument amusing. The immediate point, however, is this: we have

[68] See §3 in "On the impossible pilot wave," which appears as Essay 17 in the collection cited previously.[66] Also the appendix to the paper reprinted as Essay 10: "Einstein–Podolsky–Rosen experiments."

shown (104) to originate in (106); we have, in other words, shown that (117.1) would follow from

$$\left.\begin{aligned} P(\boldsymbol{x},\boldsymbol{p},t) &= \mathsf{R}^2(\boldsymbol{x},t)\cdot\delta(\boldsymbol{p}-\boldsymbol{\nabla}\mathsf{S}) \\ &= (\text{probability density})\cdot\delta\Big(\boldsymbol{p}-m\,\frac{\text{probability current}}{\text{probability density}}\Big) \\ &= \psi^*\psi\cdot\delta\Big(\boldsymbol{p}-\tfrac{1}{2}i\hbar\,\frac{\psi\boldsymbol{\nabla}\psi^*-\psi^*\boldsymbol{\nabla}\psi}{\psi^*\psi}\Big) \\ &= \psi^*\psi\cdot\delta\Big(\boldsymbol{p}-\tfrac{1}{2}i\hbar\big[\boldsymbol{\nabla}\log\psi^*-\boldsymbol{\nabla}\log\psi\big]\Big) \end{aligned}\right\} \tag{121}$$

which provide alternative formulations of Bohm's fundamental assumption: he takes the "pilot wave" ψ to be simultaneously responsible for

- the instantaneous design of the surface Σ_t, and
- how Σ_t is "populated"

with the result that in his theory "motion depends upon the population" (through, he would say, the "quantum potential"). There are, however, other ways than (121) to get from $\psi(\boldsymbol{x})$ into phase space. One of those—fundamental to the so-called "phase space formulation of quantum mechanics"—has already been mentioned;[40] I refer to the process

$$\psi(x) \longrightarrow P(x,p) = \tfrac{2}{\hbar}\int \psi^*(x+\xi)e^{2\frac{i}{\hbar}p\xi}\psi(x-\xi)\,d\xi \tag{122}$$

devised by E. P. Wigner and L. Szilard sometime prior to 1932. The Bohm distribution (121) and the Wigner distribution (122), if extracted from identical ψ functions, yield identical marginal distributions

$$\int P(x,p)\,dp = |\psi(x)|^2$$

which, in their separate ways, they launch identically into motion.[69] But when plotted, $P_{\text{Wigner}}(x,p)$ and $P_{\text{Bohm}}(x,p)$ could hardly be more different; the former (if we make allowances for the fact that it can assume negative values) resembles Figure 9^{upper}, while the latter (except for the fact that it can never display multivaluedness) resembles Figure 9_{lower}. Holland,[70] after remarking that "there have been many attempts to introduce phase-space-type structures into quantum mechanics," discusses only one—the formalism based upon the Wigner distribution, which he promptly dismisses on the grounds that it does not conform to Bohm's conception of the world ("does not appear to provide a suitable language for a causal representation of quantum phenomena"). The "phase space formulation of quantum mechanics," as elaborated by J.E. Moyal,[71]

[69] On this evidence we conclude that while probability conservation (91) permits, it *does not enforce* Bohm's structural assumption (121).

[70] See §8.4.3 in the monograph[61] cited previously.

[71] "Quantum mechanics as a statistical theory," Proc. Camb. Phil. Soc. **45**, 99 (1949).

is, in fact, *precisely equivalent* to orthodox quantum mechanics, of which it is an informative reorganization, but into which it imports no fundamentally new physical concepts (no gratuitous innovations subject defense/refutation only by philosophers); therein, for many, lies one of its strengths.[72]

The role of the Schrödinger equation is (in the one-dimensional case) taken over within the phase space formalism by the equation[73]

$$\tfrac{\partial}{\partial t}P = \tfrac{2}{\hbar}\sin\Big\{\tfrac{\hbar}{2}\big[\big(\tfrac{\partial}{\partial x}\big)_H\big(\tfrac{\partial}{\partial p}\big)_P - \big(\tfrac{\partial}{\partial x}\big)_P\big(\tfrac{\partial}{\partial p}\big)_H\big)\big]\Big\}HP \tag{123.1}$$

which gives back Liouville's equation (105) in the limit $\hbar \downarrow 0$. Expanding the sine, we find (123.1) to be a partial differential equation of *infinite order* (though in special cases the power series may truncate). The question which I take this opportunity to pose—but will not attempt to resolve—is this: Can (123.1) be rendered into the language of Lagrangian field theory? The question acquires urgency for the same reason as motivated us when we looked at the beginning of this discussion to the Hamilton-Jacobi equation: the "quantum Liouville equation" (123.1) can plausibly be claimed to be "fundamental." And it acquires formal interest from the circumstance that if a suitable Lagrangian could be found, it would have necessarily the form

$$\mathfrak{L}(P, \partial P, \partial\partial P, \partial\partial\partial P, \cdots)$$

The field equation (123.1) can alternatively be cast as an integral equation

$$\tfrac{\partial}{\partial t}P(x,p,t) = \iint \mathcal{K}(x,p;x_0,p_0)P(x_0,p_0,t)\,dx_0dp_0 \tag{123.2}$$

$$\mathcal{K}(x,p;x_0,p_0) \text{ suitably defined}^{73}$$

which speaks on its face of *non-local* field theory. Whether we were to proceed from (123.1) or from (123.2), we would be obligated to undertake at the outset a major *enlargement of Lagrangian field theory*, as I have presented it.

---

[72] For an excellent brief account of the phase space formalism, see pp. 422–425 in Jammer.[65]

[73] For detailed discussion see QUANTUM MECHANICS (1967), Chapter 3, p.110.

2

Introduction to the Principles of

RELATIVISTIC FIELD THEORY

Introduction. We have several times taken passing notice of what appeared to be a natural "relativistic predisposition" of classical field theory—of field theories in general. The reason for this state of affairs is not far to find: it was clearly articulated more than ninety years ago by Hermann Minkowski, who in 1908 had occasion to speak as follows:[1]

> *"The views of space and time which I wish to lay before you have sprung from the soil of experimental physics, and therein lies their strength. They are radical. Henceforth, space by itself, and time by itself, are doomed to fade away into mere shadows, and only a kind of union of the two will preserve an independent reality."*

In the dynamics of particles, as formulated by Newton and carried to a kind of perfection by Lagrange/Hamilton, the space/time distinction remains vivid; one is concerned in that theory with expressions $\boldsymbol{x}(t)$ into which $\boldsymbol{x}$ enters as the dependent variable, t (recorded by "the clock on God's wall") as the independent variable. In field theories, on the other hand, one is concerned with expressions of the form $\varphi(\boldsymbol{x}, t)$; the spatial variables $\boldsymbol{x}$ have joined the temporal variable t in an *expanded list of conceptually distinct but formally co-equal independent variables*. "Spacetime" has come into being as the 4-dimensional plenum upon which field theories are written.

[1] I quote from the introduction to his "Space and time," which is the text of an address delivered on 21 September 1908 before the 80$^{\text{th}}$ Assembly of German Natural Scientists and Physicians, in Cologne. The paper, in English translation, can be found *The Principle of Relativity* (1923), which is available as a Dover reprint.

This is not to say that field theory is "automatically relativistic." It remains to turn the "plenum" into a specifically structured metric manifold, as also it remains to stipulate that inertial observers are interconnected by transformations which preserve that metric structure. But field theory is predisposed to favor such developments. . . and historically it was a field theory (electrodynamics) which stimulated those developments—developments which led to articulation of the Principle of Relativity.

To phrase the point another way: Hamilton's principle, as encountered in particle mechanics

$$\delta \int_{\text{time interval}} L\, dt = 0$$

is anti-relativistic in that it assigns a preferred place to the temporal variable t. But its field-theoretic counterpart

$$\delta \iiiint_{\text{spacetime bubble}} \mathfrak{L}\, dxdydzdt = 0$$

is "pro-relativistic" in the obvious sense that it assigns formally identical roles to each of the spacetime coordinates.

So diverse are the distributed systems encountered in Nature that field theory has an unruly tendency to sprawl. Even after discarding all aspects of the topic (some physically quite important) which fall outside the ruberic of "Lagrangian field theory," one is left with potential subject matter far too vast to be surveyed in thirty-six lectures.[2] On a previous occasion, seeking to further condense the subject, I chose to treat only relativistic fields,[3] but found even that restricted topic to be much too broad (and in many respects too advanced) for comprehensive treatment in such a setting. Here I propose to examine only some introductory aspects of relativistic classical field theory. The specific systems we will be discussing have been selected to expose characteristic points of principle and methodology, and to provide the basic stock of concrete examples upon which we will draw in later work.

Notational conventions & relativistic preliminaries. Honoring an almost universal convention, we will use Greek indices to distinguish spacetime coordinates x^μ. Specifically

$$\begin{aligned} x^0 &= ct \\ x^1 &= x \\ x^2 &= y \\ x^3 &= z \end{aligned}$$

[2] On can, in this light, understand why L. D. Landau & E. M. Lifshitz, writing under the title *The Classical Theory of Fields* (2nd edition 1962), elect actually to treat only two fields: the electromagnetic field and the gravitational field.

[3] RELATIVISTIC CLASSICAL FIELDS (1973).

Spacetime acquires its Minkowskian metric structure from the metric tensor $g_{\mu\nu}$, which is taken to be given (in all frames) by[4]

$$\boldsymbol{g} \equiv \|g_{\mu\nu}\| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1}$$

Elements of the matrix $\boldsymbol{g}^{-1}$ inverse to $\boldsymbol{g}$ are denoted $g^{\mu\nu}$; thus $g^{\mu\alpha}g_{\alpha\nu} = \delta^{\mu}{}_{\nu}$ where the Einstein summation convention $\sum_{\alpha}$ is (as always) understood. We use $g^{\mu\nu}$ and $g_{\mu\nu}$ to raise and lower indices: thus $x_{\mu} = g_{\mu\alpha}x^{\alpha}$, etc. The Lorentzian inner product of a pair of 4-vectors x and y is defined/denoted

$$(x, y) \equiv x^{\mathsf{T}}\boldsymbol{g}\, y = x^{\alpha}g_{\alpha\beta}y^{\beta} = x^{\alpha}y_{\alpha} = x^0y^0 - x^1y^1 - x^2y^2 - x^3y^3$$

Special relativity contemplates the response of physical theory to *linear* transformations $x^{\mu} \longrightarrow X^{\mu} = \Lambda^{\mu}{}_{\alpha}x^{\alpha}$ which *preserve the postulated metric structure of spacetime*. Linearity entails that the elements of the transformation matrix $M^{\mu}{}_{\nu} \equiv \partial X^{\mu}/\partial x_{\nu} = \Lambda^{\mu}{}_{\nu}$ are x-independent *constants*, and therefore that contravariant vectors transform *like coordinates*:

$$\begin{aligned} V^{\mu} &= \tfrac{\partial X^{\mu}}{\partial x^{\alpha}} v^{\alpha} \\ &= \Lambda^{\mu}{}_{\alpha}v^{\alpha} \end{aligned} \tag{2}$$

This circumstance introduces an enormous formal simplification into relativity. "Preservation of metric structure" is interpreted to mean that

$$g_{\mu\nu} \longrightarrow G_{\mu\nu} = \Lambda^{\alpha}{}_{\mu}\Lambda^{\beta}{}_{\nu}g_{\alpha\beta} \quad \text{gives } g_{\mu\nu} \text{ back again:}$$

$$\boldsymbol{\Lambda}^{\mathsf{T}}\boldsymbol{g}\,\boldsymbol{\Lambda} = \boldsymbol{g} \tag{3}$$

Every inertial observer, when asked to write out the metric tensor, writes the same thing; $g_{\mu\nu}$ has become a *universally available shared commodity* (like the Kronecker tensor $\delta^{\mu}{}_{\nu}$, and like the Levi-Civita tensor density $\epsilon_{\kappa\lambda\mu\nu}$). From this central fact (which serves to distinguish the *Lorentz transformations* from all other linear transformations) follows the Lorentz-invariance of all inner products:

$$\begin{aligned} (x, y) &= \boldsymbol{x}^{\mathsf{T}}\boldsymbol{g}\,\boldsymbol{y} \\ &= \boldsymbol{x}^{\mathsf{T}}\boldsymbol{\Lambda}^{\mathsf{T}}\boldsymbol{g}\,\boldsymbol{\Lambda}\,\boldsymbol{y} \\ &= \boldsymbol{X}^{\mathsf{T}}\boldsymbol{g}\,\boldsymbol{Y} \equiv (\boldsymbol{X}, \boldsymbol{Y}) \end{aligned} \tag{4}$$

[4] It has been my practice for more than forty years to use doublestroke characters to distinguish matrices from other kinds of mathematical objects; thus $\mathbb{M} = \|M_{ij}\|$. But TEX provides a very limited set of such characters. Forced to abandon my former convention, I will here use **boldface** to accomplish that distinction.

In the case $y = x$ we obtain the Lorentz invariance of

$$s^2 \equiv (x, x) \quad : \quad \text{the Lorentzian "squared length" of } x$$

We are by this point in touch with the hyperbolic geometry of spacetime—in touch, that is to say, with the standard stuff of textbook relativity (time dilation, length contraction, breakdown of distant simultaneity and all the rest).

Turning now from spacetime to the fields inscribed upon spacetime... the phrase *scalar field* refers to a single-component field which transforms by the rule

$$\Phi(X(x)) = \varphi(x) \tag{5.0}$$

Vector fields have 4 components, and transform by the rule

$$\Phi^\mu(X(x)) = \Lambda^\mu{}_\alpha \varphi^\alpha(x) \tag{5.1}$$

Similarly, *tensor fields* of second rank transform

$$\Phi^{\mu\nu}(X(x)) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta \varphi^{\alpha\beta}(x) \tag{5.2}$$

So it goes.

Familiarly, the electromagnetic field transforms as an *antisymmetric* tensor field of second rank:

$$F^{\mu\nu}(X(x)) = \Lambda^\mu{}_\alpha \Lambda^\nu{}_\beta f^{\alpha\beta}(x) \tag{6}$$
$$f^{\alpha\beta} = -f^{\beta\alpha} \quad \Rightarrow \quad F^{\mu\nu} = -F^{\nu\mu}$$

We are used to deploying the elements of the field tensor as elements of an antisymmetric matrix[5]

$$\boldsymbol{f} \equiv \begin{pmatrix} 0 & - & - & - \\ f^{10} & 0 & - & f^{13} \\ f^{20} & f^{21} & 0 & - \\ f^{30} & - & f^{32} & 0 \end{pmatrix}$$

but if the independently specifiable elements are (in some arbitrary order[6]) deployed as a "6-vector" then (6) becomes

$$\begin{pmatrix} F^{10} \\ F^{20} \\ F^{30} \\ F^{32} \\ F^{13} \\ F^{21} \end{pmatrix} = \begin{pmatrix} M^{10}{}_{10} & M^{10}{}_{20} & M^{10}{}_{30} & M^{10}{}_{32} & M^{10}{}_{13} & M^{10}{}_{21} \\ M^{20}{}_{10} & M^{20}{}_{20} & M^{20}{}_{30} & M^{20}{}_{32} & M^{20}{}_{13} & M^{20}{}_{21} \\ M^{30}{}_{10} & M^{30}{}_{20} & M^{30}{}_{30} & M^{30}{}_{32} & M^{30}{}_{13} & M^{30}{}_{21} \\ M^{32}{}_{10} & M^{32}{}_{20} & M^{32}{}_{30} & M^{32}{}_{32} & M^{32}{}_{13} & M^{32}{}_{21} \\ M^{13}{}_{10} & M^{13}{}_{20} & M^{13}{}_{30} & M^{13}{}_{32} & M^{13}{}_{13} & M^{13}{}_{21} \\ M^{21}{}_{10} & M^{21}{}_{20} & M^{21}{}_{30} & M^{21}{}_{32} & M^{21}{}_{13} & M^{21}{}_{21} \end{pmatrix} \begin{pmatrix} f^{10} \\ f^{20} \\ f^{30} \\ f^{32} \\ f^{13} \\ f^{21} \end{pmatrix}$$

---

[5] See, for example, my CLASSICAL ELECTRODYNAMICS (1980), p. 162. The components $\{f^{10}, f^{20}, f^{30}\}$ are familiar as the components of the $\vec{E}$ field, and $\{f^{32}, f^{13}, f^{21}\}$ as components of the $\vec{B}$ field

[6] I select the order which in another notation reads $\{E_1, E_2, E_3, B_1, B_2, B_3\}$.

with $M^{\mu\nu}{}_{\alpha\beta} \equiv \Lambda^{\mu}{}_{\alpha}\Lambda^{\nu}{}_{\beta} - \Lambda^{\mu}{}_{\beta}\Lambda^{\nu}{}_{\alpha}$. And this invites the notational simplification

$$F^a = M^a{}_b(\mathbf{\Lambda}) f^b \quad : \quad a \ \& \ b \text{ range on } \big\{1,2,3,4,5,6\big\} \tag{7}$$

We expect to be able to show that (7) is "transformationally stable," in the sense that

$$M^a{}_p(\mathbf{\Lambda}')M^p{}_b(\mathbf{\Lambda}'') = M^a{}_b(\mathbf{\Lambda}''\mathbf{\Lambda}') \tag{8}$$

Which is to say: we expect (7) to provide a 6×6 *matrix representation of the Lorentz group*. We have stepped here unwittingly onto the shore of a vast mathematical continent—"group representation theory," as in all of its intricate parts it pertains to the Lorentz group.[7]

In brief continuation of the preceding discussion: Electrodynamical theory assigns major importance to the "dual" of the field tensor $f^{\mu\nu}$, which is defined/ denoted

$$f^{\star\mu\nu} = g^{\mu\alpha}g^{\nu\beta}f^{\star}{}_{\alpha\beta}$$
$$f^{\star}{}_{\alpha\beta} \equiv \tfrac{1}{2}\epsilon_{\alpha\beta\rho\sigma}f^{\rho\sigma}$$

The effect of "dualization" $f^{\mu\nu} \to f^{\star\mu\nu}$ can in $\big\{\vec{E},\vec{B}\big\}$-notation be described

$$\begin{pmatrix} 0 & -E_1 & -E_2 & -E_3 \\ E_1 & 0 & -B_3 & B_2 \\ E_2 & B_3 & 0 & -B_1 \\ E_2 & -B_2 & B_1 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & B_1 & B_2 & B_3 \\ -B_1 & 0 & -E_3 & E_2 \\ -B_2 & E_3 & 0 & -E_1 \\ -B_2 & -E_2 & E_1 & 0 \end{pmatrix}$$

of which $\big\{\vec{E},\vec{B}\big\} \to \big\{-\vec{B},\vec{E}\big\}$ serves to capture the essence. Moreover, one has[8]

$$\tfrac{1}{2}f^{\alpha\beta}f_{\beta\alpha} = -\tfrac{1}{2}f^{\star\alpha\beta}f^{\star}{}_{\beta\alpha} = \vec{E}\cdot\vec{E} - \vec{B}\cdot\vec{B}$$
$$-\tfrac{1}{2}f^{\ \alpha\beta}f^{\star}{}_{\beta\alpha} = \vec{E}\cdot\vec{B} + \vec{B}\cdot\vec{E} = 2\vec{E}\cdot\vec{B}$$

---

[7] To gain a preliminary sense of the landscape, see (to select but three titles from a very long shelf) S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (1961), Chapter2; J. F. Cornwell, *Group Theory in Physics* (1984), Chapter 17; F. D. Murnaghan, *The Theory of Group Representations*,(1938), Chapter 12. Classification of the unitary representations was accomplished in this pair of classic papers: E. P. Wigner, "On unitary representations of the inhomogeneous Lorentz group," Ann. of Math. **40**, 149 (1939); V. Bargmann & E. P. Wigner, "Group theoretical discussion of relativistic wave equations," PNAS **34**, 211 (1948). In the latter connection, see also Iv. M. Shirakov, "A group-theoretical consideration of the basis of relativistic quantum mechanics. III. Irreducible representations. . . of the inhomogeneous Lorentz group," Soviet Physics JETP **6**, 929 (1958). But non-unitary representations are also of great physical importance.

[8] See pp. 256 & 298 in the notes previously cited.[5] Levi-Civita dualization is treated on pp. 163–166.

—the Lorentz invariance of which is in each case manifest. Within the 6-vector formalism we are led therefore to write

$$f = \begin{pmatrix} \vec{E} \\ \vec{B} \end{pmatrix} \rightarrow f^\star = \begin{pmatrix} -\vec{B} \\ \vec{E} \end{pmatrix}$$

and to note that we are placed thus in position to write

$$\left.\begin{aligned} \vec{E}\cdot\vec{E} - \vec{B}\cdot\vec{B} &= f^a G_{ab} f^{\ b} \\ -\vec{E}\cdot\vec{B} - \vec{B}\cdot\vec{E} &= f^a G_{ab} f^{\star b} \end{aligned}\right\} \tag{9}$$

provided we set

$$\boldsymbol{G} \equiv \|G_{ab}\| = \begin{pmatrix} +1 & 0 & 0 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & +1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

The Lorentz invariance of (9) would then follow from

$$\boldsymbol{M}^{\intercal}\boldsymbol{G}\boldsymbol{M} = \boldsymbol{G} \tag{10}$$

We may consider $\boldsymbol{G}$ to deposit "induced metric structure" on 6-space, and note that (10) is formally identical to (3). We recall finally that for electromagnetic radiation it is the case that $\vec{E}\cdot\vec{E} = \vec{B}\cdot\vec{B}$, which entails that f be "null" in the sense that $(f, f) \equiv f^a G_{ab} f^b = 0$. Returning now to the main line of this discussion...

If $\varphi(x)$ transforms as a (weightless) scalar field (which is to say: by the rule (5.0)) then

$$\partial\Phi/\partial X^\mu = \sum_\alpha (\partial x^\alpha/\partial X^\mu)(\partial\varphi/\partial x^\alpha)$$

which we express

$$\Phi_{,\mu} = \tfrac{\partial x^\alpha}{\partial X^\mu}\, \varphi_{,\alpha}$$

In words: the first partials of a scalar field transform tensorially, as components of a (weightless) covariant *vector* field.

A second differentiation gives

$$\partial^2\Phi/\partial X^\mu \partial X^\nu = \sum_{\alpha\beta} (\partial x^\alpha/\partial X^\mu)(\partial x^\beta/\partial X^\nu)(\partial^2\varphi/\partial x^\alpha \partial x^\beta) + \sum_\alpha (\partial^2 x^\alpha/\partial X^\mu \partial X^\nu)(\partial\varphi/\partial x^\alpha)$$

Evidently we can write

$$\Phi_{,\mu\nu} = \tfrac{\partial x^\alpha}{\partial X^\mu}\tfrac{\partial x^\beta}{\partial X^\nu}\, \varphi_{,\alpha\beta}$$

and say that "the set of second partials transforms tensorially" only under conditions so special as to insure that the term $\sum_\alpha$(etc.) vanishes. And that it certainly does when the functions $x^\mu(X)$ depend only *linearly*—which is to say: *at most* linearly—upon their arguments... which in special relativity (inertial *Cartesian* frames interrelated by *Lorentz* transformations) is precisely the case. We conclude that, within the setting afforded by relativity, ordinary differentiation

$$\partial_\mu \text{ sends } (\text{tensor})^{\text{contravariant rank}}_{\text{covariant rank}} \longrightarrow (\text{tensor})^{\text{contravariant rank}}_{\text{covariant rank}+1}$$

This follows even if we extend the set of allowed transformations to include "inhomogeneous Lorentz transformations"[9]

$$x \to X = \mathbf{\Lambda} x + a$$

One fussy detail remains to be considered: if

$$X^\mu = M^\mu{}_\alpha x^\alpha$$

then

$$x^\alpha = W^\alpha{}_\nu X^\nu \quad \text{with} \quad \boldsymbol{W} = \|W^\alpha{}_\nu\| \equiv \boldsymbol{M}^{-1}$$

and to describe the transform rule for a mixed tensor we would write something like

$$\Phi^{\lambda\mu}{}_\nu = M^\lambda{}_\alpha M^\mu{}_\beta W^\gamma{}_\nu \, \varphi^{\alpha\beta}{}_\gamma$$

To underscore a presumption that $x \to X$ by *Lorentz* transformation we might specialize the notation, writing

$$X^\mu = \Lambda^\mu{}_\alpha x^\alpha$$
$$x^\alpha = V^\alpha{}_\nu X^\nu \quad \text{with} \quad \boldsymbol{V} = \|V^\alpha{}_\nu\| \equiv \mathbf{\Lambda}^{-1}$$

and $\Phi^{\lambda\mu}{}_\nu = \Lambda^\lambda{}_\alpha \Lambda^\mu{}_\beta V^\gamma{}_\nu \, \varphi^{\alpha\beta}{}_\gamma$. But in relativity we know something special about the construction of $\boldsymbol{V}$, for (3) supplies

$$\boldsymbol{V} \equiv \mathbf{\Lambda}^{-1} = \boldsymbol{g}^{-1}\mathbf{\Lambda}^{\mathsf{T}}\boldsymbol{g} = \|g^{\mu\alpha}\Lambda^\beta{}_\alpha g_{\beta\nu}\| = \|\Lambda_\nu{}^\mu\|$$

We can, on this basis, drop V from our list of busy symbols: the statement $\mathbf{\Lambda}\mathbf{\Lambda}^{-1} = \boldsymbol{I}$ becomes $\Lambda^\mu{}_\alpha \Lambda_\nu{}^\alpha = \delta^\mu{}_\nu$, and to describe the Lorentz transform our our mixed tensor we write

$$\Phi^{\lambda\mu}{}_\nu = \Lambda^\lambda{}_\alpha \Lambda^\mu{}_\beta \Lambda_\nu{}^\gamma \, \varphi^{\alpha\beta}{}_\gamma$$

[9] Such transformations are elements of the so-called *Poincaré group*, which includes both the Lorentz group (boosts and spatial rotations) and the group of spacetime translations as subgroups.

Contraction on the last pair of indices would be accomplished

$$\begin{aligned}\Phi^{\lambda\mu}{}_{\mu} &= \Lambda^{\lambda}{}_{\alpha}\Lambda^{\mu}{}_{\beta}\Lambda_{\mu}{}^{\gamma}\,\varphi^{\alpha\beta}{}_{\gamma} \\ &= \Lambda^{\lambda}{}_{\alpha}\delta^{\gamma}{}_{\beta}\,\varphi^{\alpha\beta}{}_{\gamma} \\ &= \Lambda^{\lambda}{}_{\alpha}\,\varphi^{\alpha\beta}{}_{\beta} \quad : \quad \varphi^{\alpha\beta}{}_{\beta} \text{ transforms as a contravariant vector}\end{aligned}$$

Objects of type $\varphi^{\mu_1\mu_2\cdots\mu_r}$ possess $2^n - 1$ "siblings," which transform in a variety of ways, but are from a practical point of view virtually interchangeable, since the transformational distinctions are managed automatically by index placement; for example,

$$\varphi^{\mu\nu} \text{ has "siblings" } \begin{cases} \varphi_{\mu}{}^{\nu} = g_{\mu\alpha}\,\varphi^{\alpha\nu} \\ \varphi^{\mu}{}_{\nu} = g_{\nu\beta}\,\varphi^{\mu\beta} \\ \varphi_{\mu\nu} = g_{\mu\alpha}g_{\nu\beta}\,\varphi^{\alpha\beta} \end{cases}$$

In special relativity the situation acquires enhanced coherence from the circumstance that the metric structure of spacetime is at every point the same:

$$g_{\mu\nu} \text{ is } x\text{-independent} \quad : \quad g_{\mu\nu,\alpha} = 0 \quad (\text{all } \mu, \nu, \alpha)$$

From this it follows that the *sibling of the derivative is the derivative of the sibling*.

Objects of type $\varphi^{\mu\nu}$ (say) respond to $x \to \mathbf{\Lambda}x$ by folding amonst themselves in $4^2 = 16$-dimensional representation of the Lorentz group, though

- if we impose the symmetry condition $\varphi^{\mu\nu} = \varphi^{\nu\mu}$ (which would make tensorial good sense) then only ten components are independently specifiable, and we are led to a 10-dimensional representation;
- if we impose the antisymmetry condition $\varphi^{\mu\nu} = -\varphi^{\nu\mu}$ then only six components are independently specifiable, and we are led to a 6-dimensional representation (as previously discussed).

The point I would emphasize is this: recent discussion has been tacitly specific to the *tensor representations of the Lorentz group*. There exists, however, a second broad class of representations—the so-called *spinor* representations. Corresponding statements are, for the most part, similar or identical, though there are some important differences. I reserve discussion of that aspect of our subject until we have physical reason to consider specific examples.

Principles of Lagrangian construction. The Lagrangian density $\mathfrak{L}(\varphi, \partial\varphi, x)$ came first to our attention as a number-valued participant in the equation

$$S_{\mathcal{R}}[\varphi] = \tfrac{1}{c}\int_{\mathcal{R}} \mathfrak{L}(\varphi, \partial\varphi, x)\; dx^0 dx^1 dx^2 dx^3 \tag{1–18}$$

which serves to define the field-theoretic action functional. Insofar as

- complex-valued fields sometimes command our physical attention, and
- $\mathfrak{L}$ sometimes enters nakedly into expressions of direct physical significance, expressions[10] which would become uninterpretable if $\mathfrak{L}$ itself were complex

it becomes pertinent to stipulate that $\mathfrak{L}(\varphi, \partial\varphi, x)$ will be required to be a *real*-valued function of its arguments.[11] I suspect that one could, on independent grounds, argue that $S_{\mathcal{R}}[\varphi]$ must, of necessity, be real (recall that analytic functions assume extreme values only at boundary points), and that the reality of $\mathfrak{L}(\varphi, \partial\varphi, x)$ is on those grounds forced, but I do not at present know how to develop the details of such an argument.

Elementary multi-variable calculus supplies the information that integrals respond to changes of variable $x \to y = y(x)$ by the rule

$$\begin{aligned}\iint\cdots\int_{\text{bubble}} & f(x)\, dx^1 dx^2 \cdots dx^n \\ &= \iint\cdots\int_{\text{image bubble}} \underbrace{f(x(y)) \left|\frac{\partial(x^1, x^2, \ldots, x^n)}{\partial(y^1, y^2, \ldots, y^n)}\right|}_{F(y)}\, dy^1 dy^2 \cdots dy^n\end{aligned}$$

The implication is that integrands transform not as simple scalars but as *scalar densities of weight* $\omega = 1$,[12] and that so, in particular, does $\mathfrak{L}$. But in the context afforded by special relativity it follows from (3) that

$$\left|\frac{\partial(\ x^0,\ x^1,\ x^2,\ x^3)}{\partial(X^0, X^1, X^2, X^3)}\right| = \left|\frac{\partial(X^0, X^1, X^2, X^3)}{\partial(\ x^0,\ x^1,\ x^2,\ x^3)}\right|^{-1} = (\det \mathbf{\Lambda})^{-1} = \pm 1$$

So to the considerable extent that we can restrict our attention to "proper" Lorentz transformations (i.e., to those which are continuous with the identity, and therefore have $\det \mathbf{\Lambda} = +1$) the tensor/tensor density distinction becomes a "distinction without a difference." We will agree to file the point away as a subtlty, to be recalled when it makes a difference.

To achieve automatic Lorentz covariance (i.e., to achieve form-invariance of the field equations, and of all that follows from them) we require that $x \to \mathbf{\Lambda} x$

---

[10] The description (1–41) of the energy density provides an example.

[11] Look again, in this light, to (1–76).

[12] For discussion of the "tensor density" concept see, for example, p. 172 of the notes previously cited.[5] We have returned here to what historically was, in fact, the birthplace of that concept. Recall in this connection that it is only by virtue of the presumption that it transforms as a density (of weight $\omega = -1$) that the Levi-Civita tensor acquires the property that

$$\epsilon_{i_1 i_2 \ldots i_n} = \text{sgn}\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

holds *in all coordinate systems*.

sends

$$\mathfrak{L}(\varphi, \partial\varphi, x) \to \textit{same} \text{ function of the } \textit{transformed} \text{ arguments}$$

And to achieve automatic implementation of that idea, we require that $\mathfrak{L}$ be expressible as a real-valued function of the *invariants* $I_1, I_2, \ldots$ which can be assembled from

- the fields $\varphi(x)$ supplied by the system $\mathcal{S}$ under consideration;
- the first partials $\partial\varphi(x)$ of those fields;
- the universally available objects $g_{\mu\nu}$ and $\epsilon_{\mu\nu\rho\sigma}$

—to which list we might, in exceptional cases, also adjoin the raw spacetime coordinates x^μ. The force of the program just sketched is made most vividly evident by concrete example. Look, therefore, to the field system $\mathcal{S}$ in which the players are a vector field A_ν and a tensor field $B_{\rho\sigma}$; by examination—tinkering: I know of no more systematic procedure, though doubtless one could be devised (the problem is combinatorial/graph-theoretic)—we are led to

$$\left.\begin{aligned} I_1 &= \partial_\alpha A^\alpha \\ I_2 &= B_\alpha{}^\alpha \end{aligned}\right\} \text{ linear in the fields}$$

$$\left.\begin{aligned} I_3 &= A_\alpha A^\alpha \\ I_4 &= (\partial^\alpha A^\beta)(\partial_\alpha A_\beta) \\ I_5 &= (\partial^\alpha A^\beta)(\partial_\beta A_\alpha) \\ I_6 &= \epsilon_{\alpha\beta\rho\sigma}(\partial^\alpha A^\beta)(\partial^\rho A^\sigma) \\ I_7 &= (\partial_\alpha A_\beta)B^{\alpha\beta} \\ I_8 &= (\partial_\alpha A_\beta)B^{\beta\alpha} \\ I_9 &= A_\alpha(\partial_\beta B^{\alpha\beta}) \\ I_{10} &= A_\alpha(\partial_\beta B^{\beta\alpha}) \\ I_{11} &= (\partial_\alpha B^{\alpha\rho})(\partial_\beta B^\beta{}_\rho) \\ I_{12} &= (\partial_\alpha B^{\alpha\rho})(\partial_\beta B_\rho{}^\beta) \\ I_{13} &= (\partial_\alpha B^{\rho\alpha})(\partial_\beta B_\rho{}^\beta) \\ I_{14} &= (\partial_\alpha B^{\alpha\rho})(\partial_\rho B_\beta{}^\beta) \\ I_{15} &= (\partial_\alpha B^{\rho\alpha})(\partial_\rho B_\beta{}^\beta) \\ I_{16} &= (\partial_\rho B^{\alpha\beta})(\partial^\rho B_{\alpha\beta}) \\ I_{17} &= (\partial_\rho B^{\alpha\beta})(\partial^\rho B_{\beta\alpha}) \\ I_{18} &= \epsilon_{\alpha\beta\rho\sigma}(\partial^\alpha A^\beta)(\partial^\rho A^\sigma) \\ I_{19} &= \epsilon_{\alpha\beta\rho\sigma}B^{\alpha\beta}B^{\rho\sigma} \\ I_{20} &= \epsilon_{\alpha\beta\rho\sigma}(\partial^\alpha A^\beta)B^{\rho\sigma} \\ I_{21} &= \epsilon_{\alpha\beta\rho\sigma}A^\alpha(\partial^\beta B^{\rho\sigma}) \end{aligned}\right\} \text{ quadratic in the fields}$$

$$\left.\begin{aligned} I_{22} &= A_\alpha A_\beta B^{\alpha\beta} \end{aligned}\right\} \text{ cubic in the fields}$$

I may have missed some, but you get the idea. Higher order invariants factor into products of those listed. The list would be shortened if (anti)symmetry properties were imposed upon $B_{\rho\sigma}$, much lengthened if either the rank or the number of the participating fields were increased. It would be lengthened also if we allowed ourselves to introduce translational symmetry-breaking terms such as $I_{21} = x^\alpha A_\alpha$.

Relativistically covariant accounts of the dynamics of such systems would now result from writing

$$\mathcal{L}(A, B, \partial A, \partial B) = \ell(I_1, I_2, \ldots, I_{22})$$

But what functional form should be assigned to ℓ(etc.)? As physicists it has been our reductionistic practice—a practice which has enjoyed a high degree of success, but which is supported no physically/philosophically secure foundation, and which in this wholistic world may be doomed to ultimate failure—to resolve systems into their imagined "component parts," and then to study how those parts interact. It becomes in this light natural to write

$$\mathcal{L}(A, B, \partial A, \partial B) = \mathcal{L}_A(A, \partial A) + \mathcal{L}_B(B, \partial B) + g \cdot \mathcal{L}_{\text{interaction}}(A, B, \partial A, \partial B)$$

where $\mathcal{L}_A(A, \partial A)$ describes the dynamics of "free A-fields," $\mathcal{L}_B(B, \partial B)$ the dynamics of "free B-fields," and the control parameter g describes the strength of the (typically weak) interaction. But we are brought thus only part way to the resolution of our problem, for...

When we look (say) to $\mathcal{L}_A(A, \partial A) = \ell_A(I_1, I_3, I_4)$ we still confront this question: What functional structure should we assign to ℓ_A(etc.)? Here major simplification results, and some important physics is brought into view, if we declare a special interest in *linear* field theories—theories dominated by a *principle of superposition*. For then $\mathcal{L}_A(A, \partial A)$ has necessarily to be *quadratic* in its arguments,[13] and the number of available options is greatly reduced: we have

$$\begin{aligned}\mathcal{L}_A(A, \partial A) &= c_4 I_4 + c_5 I_5 + c_6 I_6 + c_1 I_1^2 + c_3 I_3 \\ &= c_4(\partial^\alpha A^\beta)(\partial_\alpha A_\beta) + c_5(\partial^\alpha A^\beta)(\partial_\beta A_\alpha) + c_6\, \epsilon_{\alpha\beta\rho\sigma}(\partial^\alpha A^\beta)(\partial^\rho A^\sigma) \\ &\qquad + c_1(\partial_\alpha A^\alpha)(\partial_\beta A^\beta) + c_3 A_\alpha A^\alpha\end{aligned}$$

where $c_1, c_3, \ldots, c_6$ are adjustable constants. By a similar argument

$$\mathcal{L}_B(B, \partial B) = \text{linear combination of } I_{11}, I_{12}, I_{13}, I_{14}, I_{15}, I_{16}, I_{17}, I_{19} \text{ and } I_2^2$$

[13] More formally, it has to be *homogeneous of degree 2:*

$$\mathcal{L}_A(\lambda A, \lambda \partial A) = \lambda^2 \mathcal{L}_A(A, \partial A)$$

It is with less conviction that we would assert $\mathfrak{L}_{\text{interaction}}$ to be a linear combination of $I_7, I_8, I_9, I_{10}, I_{20}$ and I_{21}, for interaction (external forcing) typically entails a suspension of the principle of superpositon. Recall the situation in electrodynamics, where

$$\partial_\mu F^{\mu\nu} = \begin{cases} 0 & \text{in the absence of sources} \\ J^\mu & \text{in the presence of sources} \end{cases}$$

Solutions of the former equation yield other solutions when added, solutions of the latter equation do not (though to any particular solution of the latter one can add arbitrary solutions of the former).

To summarize: Special relativity (acting conjointly with the principle of superposition) exerts a fairly strong constraint upon the design of free field theories, but to achieve a convincing "theory of interaction" we need the input of a new idea.[14]

One final remark before we turn to discussion of some particular free field theories: Such structural properties as we may impute to $\mathfrak{L}$ are susceptible to seeming contravention by gauge transformation

$$\mathfrak{L} \longrightarrow \mathfrak{L}' = \mathfrak{L} + \partial_\alpha \mathcal{G}^\alpha \tag{1–23}$$

For $\mathcal{G}^\alpha$—which can be "anything"—is free to violate just about any condition we may have in mind.

Real scalar field: the Klein-Gordon equation. Suppose $\mathcal{S}$ affords only a single real-valued scalar field $\varphi(x)$. The quadratic invariants which can be assembled from φ and $\partial_\mu\varphi$ are two in number:

$$I_1 = (\partial^\alpha\varphi)(\partial_\alpha\varphi) \quad \text{and} \quad I_2 = \varphi^2$$

so we are led to write

$$\begin{aligned} \mathfrak{L}(\varphi, \partial\varphi) &= c_1 I_1 + c_2 I_2 \\ &= \tfrac{1}{2}K\Big\{g^{\alpha\beta}\varphi_{,\alpha}\varphi_{,\beta} - \varkappa^2\varphi^2\Big\} \end{aligned} \tag{11}$$

where $[K] = (\text{energy/length})/[\varphi]^2$ and $[\varkappa^2] = 1/(\text{length})^2$. The resulting field equation

$$\Big\{\partial_\mu \frac{\partial}{\partial\varphi_{,\mu}} - \frac{\partial}{\partial\varphi}\Big\}\mathfrak{L} = \partial_\mu(g^{\mu\alpha}\varphi_{,\alpha}) + \varkappa^2\varphi = 0$$

[14] It was partly to achieve the latter objective that *gauge field theory* was invented; see the introduction to Chapter 9 in L. O'Raifeartaigh's *The Dawning of Gauge Theory* (1997), which alludes to the motivational pulse of Ronald Shaw's doctoral research (1952–54). Shaw's thesis joins a classic paper by C. N. Yang & R. Mills ("Isotopic spin conservation and a generalized gauge invariance," PR **95**, 631 (1954)) as one of the historic sources of gauge theory in its modern form.

can be written

$$\big(\square + \varkappa^2\big)\varphi = 0 \tag{12}$$

where

$$\square \equiv g^{\alpha\beta}\partial_\alpha\partial_\beta = \tfrac{1}{c^2}\big(\tfrac{\partial}{\partial t}\big)^2 - \nabla^2 \tag{13}$$

defines the *wave operator* (or "d'Alembertian").

At (12) we have encountered the celebrated *Klein-Gordon equation*, though the equation studied by O. Klein and W. Gordon in 1926—which had been written down but abandoned by Schrödinger himself even prior to the invention of the non-relativistic "Schrödinger equation"—was actually the complex analog of (12). Useful insight into the historical origins of (12) is gained when one sets

$$\varphi(x) = e^{\frac{i}{\hbar}(p,x)} \quad : \quad (p,x) \equiv p_\alpha x^\alpha = Et - \boldsymbol{p}\cdot\boldsymbol{x}$$

and finds that (12) will be satisfied if and only if

$$g^{\alpha\beta}p_\alpha p_\beta \equiv (E/c)^2 - \boldsymbol{p}\cdot\boldsymbol{p} = \hbar^2\varkappa^2$$

But the relativistic theory of a mass point m presents us with the "dispersion relation" $(E/c)^2 - \boldsymbol{p}\cdot\boldsymbol{p} = (mc)^2$. It becomes in this light natural to set

$$\varkappa = \frac{mc}{\hbar} \tag{14}$$

and to consider the Klein-Gordon equation (12) to be fruit of the "Schrödinger quantization" procedure

$$g^{\alpha\beta}p_\alpha p_\beta\Big|_{p\,\to\, i\hbar\partial} = (mc)^2$$

Notice that the Klein-Gordon equation gives back the wave equation $\square\varphi = 0$ in the limit $\varkappa \downarrow 0$; i.e., in the limit of zero mass.[15]

Working from (1–34) we find the stress-energy tensor of the real scalar field to be given by

$$\begin{aligned} S^\mu{}_\nu &= \frac{\partial\mathcal{L}}{\partial\varphi_{,\mu}}\,\varphi_{,\nu} - \mathcal{L}\delta^\mu{}_\nu \\ &= \tfrac{1}{2}K\Big\{(g^{\mu\alpha}+g^{\alpha\mu})\varphi_{,\alpha}\varphi_{,\nu} - (g^{\alpha\beta}\varphi_{,\alpha}\varphi_{,\beta})\delta^\mu{}_\nu + \varkappa^2\varphi^2\delta^\mu{}_\nu\Big\} \end{aligned} \tag{15}$$

[15] For more detailed discussion of the Klein-Gordon equation, see Chapter 3 in Schweber[7] or Chapter 20 §5 in A. Massiah, *Quantum Mechanics* (1966). Schrödinger abandoned his relativistic equation for the good physical reason that it led to the wrong hydrogen spectrum: see §51 in L. I. Schiff, *Quantum Mechanics* (3rd edition 1968) for details and references.

from which we learn (see again (1–41)) that the *energy density* of such a field can be described

$$\begin{aligned}\mathcal{E} = S^0{}_0 &= \tfrac{1}{2}K\Big\{2(\varphi_{,0})^2 - \big[(\varphi_{,0})^2 - (\varphi_{,1})^2 - (\varphi_{,2})^2 - (\varphi_{,3})^2\big] + \varkappa^2\varphi^2\Big\} \\ &= \tfrac{1}{2}K\Big\{(\varphi_{,0})^2 + (\varphi_{,1})^2 + (\varphi_{,2})^2 + (\varphi_{,3})^2 + \varkappa^2\varphi^2\Big\} \qquad (16.1)\end{aligned}$$

It was in anticipation of this result—i.e., to ensure the non-negativity of energy density—that we gave the name $\varkappa^2$ to the coefficient of the φ^2-term in (11). The three components of momentum density are given by

$$c\mathcal{P}_i = S^0{}_i = \frac{\partial\mathcal{L}}{\partial\varphi_{,0}}\,\varphi_{,\nu} = K\varphi_{,0}\varphi_{,i} \quad : \quad (i = 1, 2, 3) \qquad (16.2)$$

Complex scalar field. Suppose $\mathcal{S}$ affords only a single *complex*-valued scalar field

$$\begin{aligned}\psi\ (x) &= \varphi_1(x) + i\varphi_2(x) = A(x)e^{+i\phi(x)} \\ \psi^*(x) &= \varphi_1(x) - i\varphi_2(x) = A(x)e^{-i\phi(x)}\end{aligned}$$

The quadratic invariants which can be assembled from ψ, ψ^*, $\partial_\mu\psi$ and $\partial_\mu\psi^*$ are quickly listed, and lead one to contemplate Lagrange densities of the form

$$\mathcal{L} = \tfrac{1}{2}g^{\alpha\beta}\big\{a\,\psi_{,\alpha}\psi_{,\beta} + 2a_0\psi^*_{,\alpha}\psi_{,\beta} + a^*\psi^*_{,\alpha}\psi^*_{,\beta}\big\} + \tfrac{1}{2}\big\{b\,\psi\psi + 2b_0\psi^*\psi + b^*\psi^*\psi^*\big\}$$

where the postulated reality condition $\mathcal{L} = \mathcal{L}^*$ requires that a_2 and b_2 be real. If we allow ourselves to write

$$\boldsymbol{\psi} \equiv \begin{pmatrix}\psi \\ \psi^*\end{pmatrix} \quad \text{and} \quad \boldsymbol{\psi}^\dagger \equiv \begin{pmatrix}\psi^* \\ \psi\end{pmatrix}^{\mathsf{T}} = \begin{pmatrix}\psi^* & \psi\end{pmatrix}$$

and to drop the commas which signify differentiation (writing $\boldsymbol{\psi}_\mu \equiv \partial_\mu\boldsymbol{\psi}$) then the preceding $\mathcal{L}$ can be displayed in the following somewhat more orderly manner:

$$\mathcal{L} = \tfrac{1}{2}g^{\alpha\beta}\boldsymbol{\psi}^\dagger_\alpha\begin{pmatrix}a_0 & a^* \\ a & a_0\end{pmatrix}\boldsymbol{\psi}_\beta + \tfrac{1}{2}\boldsymbol{\psi}^\dagger\begin{pmatrix}b_0 & b^* \\ b & b_0\end{pmatrix}\boldsymbol{\psi}$$

The reality condition can be said in this notation to result from the *hermiticity* of the 2×2 matrices. Reality would be retained even if we were to relax the requirement that the diagonal elements be equal, but when we do so—writing

$$\mathcal{L} = \tfrac{1}{2}g^{\alpha\beta}\boldsymbol{\psi}^\dagger_\alpha\,\mathbb{A}\,\boldsymbol{\psi}_\beta + \tfrac{1}{2}\boldsymbol{\psi}^\dagger\mathbb{B}\,\boldsymbol{\psi} \qquad (17)$$

with

$$\mathbb{A} = \begin{pmatrix}a_0 + a_3 & a_1 - ia_2 \\ a_1 + ia_2 & a_0 - a_3\end{pmatrix} \quad \text{and} \quad \mathbb{B} = \begin{pmatrix}b_0 + b_3 & b_1 - ib_2 \\ b_1 + ib_2 & b_0 - b_3\end{pmatrix}$$

—it becomes apparent that the new parameters a_3 and b_3 actually make no net contribution to $\mathcal{L}$, therefore none to the field equations, none to the physics. In (17) we have, therefore, what is in effect a 6-parameter population of theories.

In an effort to sharpen the problem before us (and to prepare the ground for some future work), we—drawing inspiration from quantum mechanics—agree to look only to those theories which possess

$$\left.\begin{array}{l}\psi \;\to e^{+i\omega}\psi \\ \psi^* \to e^{-i\omega}\psi^*\end{array}\right\} \quad : \quad \text{equivalently} \quad \boldsymbol{\psi} \to \begin{pmatrix} e^{+i\omega} & 0 \\ 0 & e^{-i\omega}\end{pmatrix}\boldsymbol{\psi} \tag{18}$$

as an internal symmetry. Immediately $a = b = 0$ ($\mathbb{A}$ and $\mathbb{B}$ must, in other words, be diagonal—effectively multiples of the identity), and we a Langrangian which, after notational adjustments, can be expressed

$$\mathfrak{L} = \tfrac{1}{2}K\Big\{g^{\alpha\beta}\psi^*_{,\alpha}\psi_{,\beta} - \varkappa^2\psi^*\psi\Big\} \tag{19}$$

This, significantly, is the most general instance of our original Lagrangian which manifests the property that it is *bilinear* in starred and unstarred field variables.

The resulting field equations read

$$\left.\begin{array}{l}\Big\{\partial_\mu\dfrac{\partial}{\partial\psi_{,\mu}} - \dfrac{\partial}{\partial\psi}\;\Big\}\mathfrak{L} = \partial_\mu(g^{\mu\alpha}\psi^*_{,\alpha}) + \varkappa^2\psi^* = \big(\Box + \varkappa^2\big)\psi^* = 0 \\ \Big\{\partial_\mu\dfrac{\partial}{\partial\psi^*_{,\mu}} - \dfrac{\partial}{\partial\psi^*}\Big\}\mathfrak{L} = \partial_\mu(g^{\mu\alpha}\psi_{,\alpha}) + \varkappa^2\psi \;\; = \big(\Box + \varkappa^2\big)\psi \;\; = 0\end{array}\right\} \tag{20}$$

Bilinearity has had the consequence that

- ψ -variation yields a field equation involving only ψ^*;
- ψ^*-variation yields a field equation involving only ψ :

the equations are uncoupled; each is the complex conjugate of the other.

The stress-energy tensor of any (ψ, ψ^*)-field system can be described

$$S^\mu{}_\nu = \frac{\partial\mathfrak{L}}{\partial\psi_{,\mu}}\,\psi_{,\nu} + \frac{\partial\mathfrak{L}}{\partial\psi^*_{,\mu}}\,\psi^*_{,\nu} - \mathfrak{L}\delta^\mu{}_\nu$$

which for systems of type (19) becomes

$$S^\mu{}_\nu = \tfrac{1}{2}K\Big\{g^{\mu\alpha}\big(\psi^*_{,\alpha}\psi_{,\nu} + \psi_{,\alpha}\psi^*_{,\nu}\big) - \big(g^{\alpha\beta}\psi^*_{,\alpha}\psi_{,\beta}\big)\delta^\mu{}_\nu + \varkappa^2\psi^*\psi\,\delta^\mu{}_\nu\Big\}$$

That—in consequence ultimately of the field equations—the components of the stress-energy tensor enter into a quartet of conservation laws

$$\partial_\mu S^\mu{}_\nu = 0$$

we know already on quite general grounds; those statements follow[16] from the circumstance that x does not enter *explicitly* into the design of $\mathfrak{L}$: $\boldsymbol{\partial}_\mu\mathfrak{L} = 0$. Energy density, for systems of type (19), can be described

$$\mathcal{E} = S^0{}_0 = \tfrac{1}{2}K\Big\{\psi^*_{,0}\psi_{,0} + \psi^*_{,1}\psi_{,1} + \psi^*_{,2}\psi_{,2} + \psi^*_{,3}\psi_{,3} + \varkappa^2\psi^*\psi\Big\} \geqslant 0$$

[16] See again the discussion subsequent to (1–34).

The results just summarized derive from the translational map. For more general maps, Noether's theorem—adapted to systems of type (19)—supplies

$$J^\mu_r = \tfrac{1}{2}K\Big\{g^{\mu\alpha}\psi^*_{,\alpha}\big[\Phi_r - \psi_{,\beta}\mathcal{X}^\beta_r\big] + g^{\mu\alpha}\psi_{,\alpha}\big[\Phi^*_r - \psi^*_{,\beta}\mathcal{X}^\beta_r\big] + \big[g^{\alpha\beta}\psi^*_{,\alpha}\psi_{,\beta} - \varkappa^2\psi^*\psi\big]\mathcal{X}^\mu_r\Big\}$$

For the one-parameter internal symmetry map (18) the r-subscript can be abandoned, and we have

$$\Phi = i\psi \quad \text{and} \quad \mathcal{X}^\mu = 0$$

giving[17]

$$Q^\mu = i\,\tfrac{1}{2}Kg^{\mu\alpha}\Big\{\psi^*_{,\alpha}\psi - \psi^*\psi_{,\alpha}\Big\} \tag{21}$$

We expect to have $\partial_\mu Q^\mu = 0$, and by calculation

$$\begin{aligned}\partial_\mu Q^\mu &= i\,\tfrac{1}{2}Kg^{\mu\alpha}\Big\{\psi^*_{,\mu\alpha}\psi + \psi^*_{,\alpha}\psi_{,\mu} - \psi^*_{,\mu}\psi_{,\alpha} - \psi^*\psi_{,\mu\alpha}\Big\}\\ &= i\,\tfrac{1}{2}K\Big\{\psi\,\Box\psi^* - \psi^*\,\Box\psi\Big\} \quad \text{after a cancellation}\\ &= -i\,\tfrac{1}{2}K\Big\{\psi\,\varkappa^2\,\psi^* - \psi^*\varkappa^2\,\psi\Big\} \quad \text{by the field equations}\\ &= 0\end{aligned}$$

find this to be in fact the case *by implication of the equations of motion*.

Real vector field: Procca's equations. Suppose system $\mathcal{S}$ affords only a solitary real-valued vector field $U_\mu(x)$. The quadratic invariants which can be assembled from U and $\partial_\mu U$ are (see again the list on p. 10) five in number, and lead one to contemplate Lagrangians[18] of the form

$$\mathfrak{L} = \tfrac{1}{2}c_0U^\alpha U_\alpha + \tfrac{1}{2}\big\{c_1g^{\alpha\rho}g^{\sigma\beta} + c_2g^{\alpha\sigma}g^{\beta\rho} + c_3g^{\alpha\beta}g^{\rho\sigma} + c_4\epsilon^{\alpha\beta\rho\sigma}\big\}U_{\alpha,\beta}U_{\rho,\sigma}$$

The associated field equations can be written

$$\partial_\nu\big\{c_1U^{\mu,\nu} + c_2U^{\nu,\mu} + c_3g^{\mu\nu}U^\alpha{}_{,\alpha}\big\} - c_0U^\mu = 0$$

giving

$$c_1\Box U_\mu + (c_2 + c_3)U^\alpha{}_{,\alpha\mu} - c_0U_\mu = 0$$

[17] The notation J^μ_r is generic; typically one adopts a non-generic notation to reflect the fact that one has been led from a specialized map to an object to which one intends to give a name more specific than "Noetherean current." Here my Q^μ is intended to as an allusion to "charge."

[18] I will hereafter allow myself—in the company of the rest of the world, and except when confusion might result (which is seldom)—to say "Lagrangian" when I mean "Lagrange density."

The surprising absence of c_4 from the preceding equation can be understood as follows: define $\mathcal{G}^\sigma \equiv \epsilon^{\alpha\beta\rho\sigma}U_{\alpha,\beta}U_\rho$ and observe that

$$\partial_\sigma \mathcal{G}^\sigma = \epsilon^{\alpha\beta\rho\sigma}U_{\alpha,\beta}U_{\rho,\sigma} + \underbrace{\epsilon^{\alpha\beta\rho\sigma}U_{\alpha,\beta\sigma}U_\rho}_{\text{0 by the } \beta\sigma\text{-symmetry of } U_{\alpha,\beta\sigma}}$$

The c_4-term is absent from the equations of motion because it is a gauge term (can be "gauged away"). Evidently we lose nothing if we write

$$\mathcal{L} = \tfrac{1}{2}\{c_1 g^{\alpha\rho}g^{\sigma\beta} + c_2 g^{\alpha\sigma}g^{\beta\rho} + c_3 g^{\alpha\beta}g^{\rho\sigma}\}U_{\alpha,\beta}U_{\rho,\sigma} - \tfrac{1}{2}\varkappa^2 U^\alpha U_\alpha \tag{22}$$

which gives

$$c_1 \Box U_\mu + (c_2 + c_3)\partial_\mu(\partial_\alpha U^\alpha) + \varkappa^2 U_\mu = 0 \tag{23}$$

We have now in hand a 4-parameter population of vector field theories. How are we to isolate the most interesting specimens? Classical electrodynamics provides the vector field with which we have greatest familiarity, and it is to electrodynamics we will look for guidance. The effort will serve to lend deepened comprehension to what kind of a thing it is that Maxwell gave to the world.

The vector field of electrodynamical interest is the so-called "4-potential" $A^\mu(x)$, introduced via

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu \tag{24.1}$$

in order to render the sourceless half

$$\partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0 \tag{24.2}$$

of Maxwell's equations automatic. The sourcey other half

$$\partial_\mu F^{\mu\nu} = J^\nu \tag{24.3}$$

then become

$$\Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = J^\nu \tag{24.4}$$

from which charge conservation

$$\partial_\nu J^\nu = 0 \tag{24.5}$$

follows as an immediate corollary. Potentials A_μ and $A'_\mu = A_\mu + \partial_\mu\chi$ (with χ arbitrary) are "gauge equivalent" in the sense that both, when introducted into the right side of (24.1), yield the same $F^{\mu\nu}$. That is why the A^μ-field is called a "potential," and denied any direct claim to physicality;[19] "physicality" might be attributed to the *set* $\{A^\mu\}$ of gauge-equivalent 4-potentials, or to any

[19] The story is a familiar one; it is for an identical reason that "direct physicality" cannot be attributed to the Lagrangian $\mathcal{L}$. The spooks appear to be in control!

gauge-invariant attribute of that set, but not to its individual elements. Pick an arbitrary element A'_μ of the set, write $A_\mu = A'_\mu - \partial_\mu \chi$ and select χ to be any solution of $\Box\chi = \partial^\mu A'_\mu$; one has then achieved

$$\partial_\mu A^\mu = 0 \tag{24.6}$$

This is the "Lorentz gauge condition," which can without loss of generality always be *assumed* to pertain. When that assumption is in force one can, in place of (24.4), write the simpler equation

$$\Box A^\nu = J^\nu \tag{24.7}$$

In the absence of sources (i.e., for *free* electromagnetic fields) the right sides of (24.3/4/7) vanish, and all else remains unchanged. In particular, (23.4) becomes

$$\Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = 0 \tag{24.8}$$

Note that we have been brought thus far by Maxwell, without any reference to the resources of Lagrangian field theory.

The field equations (23) would assume the electromagnetic form (24.8) were we to set $\varkappa = 0$ and $c_2 + c_3 = -c_1$; were we, in other words, to set

$$c_1 + c_2 + c_3 = 0 \quad \text{and} \quad \varkappa = 0$$

Look now to the class of theories which results when one retains the former of those conditions but drops the latter:

$$c_1 + c_2 + c_3 = 0 \quad \text{and} \quad \varkappa \neq 0$$

The field equations (23) then read

$$c_1\big\{\Box U_\mu - \partial_\mu(\partial_\alpha U^\alpha)\big\} + \varkappa^2 U_\mu = 0 \tag{25}$$

which when hit with ∂^μ yield the "Lorentz gauge-like" statement

$$\partial_\mu U^\mu = 0 \tag{26}$$

not as an arbitrarily imposed side condition but as an *enforced corollary of the equations of motion*. Returning with this information to (25) we obtain

$$\big(\Box + \varkappa^2\big)U^\mu = 0 \tag{27}$$

The solutions of $c_1 + c_2 + c_3 = 0$ can, for given c_1, be parameterized

$$\left.\begin{aligned} c_2 &= +\lambda c_1 - c_1 \\ c_3 &= -\lambda c_1 \end{aligned}\right\} \quad : \quad \lambda \text{ arbitrary}$$

and the Lagrangian (22) in this notation becomes[20]

$$\begin{aligned}\mathfrak{L} &= \tfrac{1}{2}\{g^{\alpha\rho}g^{\sigma\beta} - (1-\lambda)g^{\alpha\sigma}g^{\beta\rho} - \lambda g^{\alpha\beta}g^{\rho\sigma}\}U_{\alpha,\beta}U_{\rho,\sigma} - \tfrac{1}{2}\varkappa^2 U^\alpha U_\alpha \\ &= \tfrac{1}{2}U^{\alpha,\beta}(U_{\alpha,\beta} - U_{\beta.\alpha}) - \tfrac{1}{2}\varkappa^2 U^\alpha U_\alpha \\ &\qquad + \lambda\underbrace{\{U^\alpha{}_{,\beta}U^\beta{}_{,\alpha} - U^\alpha{}_{,\alpha}U^\beta{}_{,\beta}\}} \\ &\qquad\qquad = \partial_\alpha\{U^\alpha{}_{,\beta}U^\beta - U^\alpha U^\beta{}_{,\beta}\}\end{aligned}$$

We might as well discard the λ-term, since it can (as indicated) always be gauged away; this done, we have

$$\mathfrak{L} = \tfrac{1}{2}U^{\alpha,\beta}(U_{\alpha,\beta} - U_{\beta,\alpha}) - \tfrac{1}{2}\varkappa^2 U^\alpha U_\alpha \tag{28.1}$$

But $U^{\alpha,\beta} = \frac{1}{2}\big(U^{\alpha,\beta} + U^{\beta,\alpha}\big) + \frac{1}{2}\big(U^{\alpha,\beta} - U^{\beta,\alpha}\big) = U^{\alpha,\beta}_{\text{symmetric}} + U^{\alpha,\beta}_{\text{antisymmetric}}$ and only the latter term survives the double summation process; it becomes therefore possible to write

$$\mathfrak{L} = \tfrac{1}{4}(U^{\alpha,\beta} - U^{\beta,\alpha})(U_{\alpha,\beta} - U_{\beta,\alpha}) - \tfrac{1}{2}\varkappa^2 U^\alpha U_\alpha \tag{28.2}$$

Proceeding in imitation of our electrodynamical experience, we introduce

$$G^{\mu\nu} \equiv \partial^\mu U^\nu - \partial^\nu U^\mu = -\big(U^{\mu,\nu} - U^{\nu,\mu}\big) \tag{29}$$

from which

$$\partial^\lambda G^{\mu\nu} + \partial^\mu G^{\nu\lambda} + \partial^\nu G^{\lambda\mu} = 0 \tag{30}$$

follows as a corollary. Notice that if (30) were—as in electrodynamics—*postulated*, then the possibility of writing (29) would be implied,[21] but U^μ would be determined *only up to gauge*, and the "Lorentz gauge-like" condition (26) would not be automatically in force. Recall also that it *is* in force only because we have assumed $\varkappa \neq 0$.

The road just travelled was first travelled by A. Procca,[22] who sought to create what was in effect a "theory of massive photons." In the following figure

[20] To simplify ensuing discussion we at this point set $c_1 = 1$; i.e., we absorb c_1 into the definitions of U^μ and $\varkappa$.

[21] Analogously: if it is known of $\boldsymbol{F}$ that $\boldsymbol{\nabla}\times\boldsymbol{F} = \boldsymbol{0}$, then $\boldsymbol{F} = -\boldsymbol{\nabla}U$ is implied, but U is determined only up to gauge: $U \to U' = U + \text{constant}$, and must therefore be considered "unphysical." But if we take U—considered "physical" (no gauge)—to be our point of departure and introduce $\boldsymbol{F} \equiv -\boldsymbol{\nabla}U$ as a *definition*, then $\boldsymbol{\nabla}\times\boldsymbol{F} = \boldsymbol{0}$ acquires the status of an "interesting corollary."

[22] "Sur la Théorie Ondulatoire des Electrons Positifs et Négatifs," Jour. Phys. Rad. **7**, 347 (1936). Procca, pursuing a fundamental idea injected into physics the previous year by H. Yukawa, sought to create a "theory of vector mesons" which adhered as closely as possible to the electromagnetic model. The subject is explored in much greated detail in my RELATIVISTIC CLASSICAL FIELDS (1973), pp. 85–97. My principal sources were W. Pauli, "Relativistic field theories of elementary particles," Rev. Mod. Phys. **13**, 203 (1941) Part II §2; E. M. Corson, *Introduction to Tensors, Spinors, and Relativistic Wave Equations* (1953), pp. 99–103; G. Wentzel, *Quantum Theory of Fields* (1949), Chapter 3.

FREE MAXWELL FIELD	FREE PROCCA FIELD
$\bullet\ \partial^\lambda F^{\mu\nu} + \partial^\mu F^{\nu\lambda} + \partial^\nu F^{\lambda\mu} = 0$	$\partial^\lambda G^{\mu\nu} + \partial^\mu G^{\nu\lambda} + \partial^\nu G^{\lambda\mu} = 0$
$\Downarrow$	$\partial_\mu G^{\mu\nu} + \varkappa^2 U^\nu = 0$
$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$	$\Uparrow$
$\bullet\quad \partial_\mu F^{\mu\nu} = 0$	$G^{\mu\nu} \equiv \partial^\mu U^\nu - \partial^\nu U^\mu$
$\Downarrow$	$\uparrow$
$\Box A^\nu = 0$ in Lorentz gauge:	$\bullet\ \big(\Box + \varkappa^2\big) U^\nu = 0$
$\partial_\mu A^\mu = 0$	$\bullet\qquad \partial_\mu U^\mu = 0$

I have attempted to display/contrast the logical pattern of the two theories in question. I have used $\bullet$ to mark the postulated free field equations, and arrows to indicate the flow of subsequent argument. Note that while in Maxwellean theory $F^{\mu\nu}$ is physical/fundamental, and A^μ an auxiliary construct, in Procca's theory it is the vector field U^μ which is fundamental, and the tensor field which plays the auxiliary role. Comparison of the two theories puts one in position to assert that *gauge transformations*

$$A_\mu \to A'_\mu = A + \partial_\mu \chi$$

enter into electromagnetic discourse in consequence of the "masslessness of the photon:" $\varkappa = 0$. In the contrary case ($\varkappa \neq 0$) the Lorentz gauge becomes not an option, but a mandate.

Procca has given us an good, unexceptionable-in-every-way example of a Lagrangian field theory. Actually a class of such examples... which have some surprising things to teach us. Notice first that if (28.2) is renotated

$$\mathcal{L} = \tfrac{1}{2}\Big\{G^{\sigma\rho}(U_{\rho,\sigma} - U_{\sigma,\rho}) - \tfrac{1}{2}G^{\rho\sigma}G_{\rho\sigma}\Big\} - \tfrac{1}{2}\varkappa^2 U^\rho U_\rho \tag{31}$$

and if U^μ and $G^{\mu\nu} = -G^{\nu\mu}$ are construed to be *independent* fields (i.e., if we suppose ourselves to be considering the interaction of a vector field with an antisymmetric tensor field), then construction of the field equations

$$\Big\{\partial_\mu \frac{\partial}{\partial G_{\alpha\beta,\mu}} - \frac{\partial}{\partial G_{\alpha\beta}}\Big\}\mathcal{L} = -\tfrac{1}{2}\Big\{\big(U^{\beta,\alpha} - U^{\alpha,\beta}\big) - G^{\alpha\beta}\Big\} = 0$$

$$\Downarrow$$

$$G^{\mu\nu} = \partial^\mu U^\nu - \partial^\nu U^\mu \tag{32.1}$$

$$\Downarrow$$

$$\therefore\quad \partial^\lambda G^{\mu\nu} + \partial^\mu G^{\nu\lambda} + \partial^\nu G^{\lambda\mu} = 0 \tag{32.2}$$

$$\Big\{\partial_\mu \frac{\partial}{\partial U_{\alpha,\mu}} - \frac{\partial}{\partial U_\alpha}\Big\}\mathcal{L} = \partial_\mu \tfrac{1}{2}\big(G^{\mu\alpha} - G^{\alpha\mu}\big) + \varkappa^2 U^\alpha = 0$$

$$\Downarrow$$

$$\partial_\mu G^{\mu\nu} + \varkappa^2 U^\nu = 0 \quad \text{by } G^{\mu\nu}\text{-antisymmetry} \tag{32.3}$$

$$\Downarrow$$

$$\therefore\quad \partial_\nu U^\nu = 0 \tag{32.4}$$

is found to yield the entire set of Procca equations (including the equation (32.1) which previously served to *define* $G^{\mu\nu}$).

We touch here on a formal problem which causes electrodynamics—historic mother of classical field theory though she is—to stand somewhat apart from the main line of field-theoretic development... as an "exceptional case." If in (32) we set $\varkappa \to 0$ then we obtain equations which differ only notationally[23] from the free-field equations of Maxwellian theory (in Lorentz gauge). But if *in the Lagrangian* (31) we set $\varkappa = 0$ then we lose the leverage which gave us (32.4); we obtain the equations of free-field electrodynamics *except for the Lorentz gauge condition*, in the absence of which the remaining equations are incorrect, but which must be brought into the theory as an unmotivated import. A. O. Barut, on p. 102 of his *Electrodynamics and Classical Theory of Fields & Particles* (1964), tabulates four different "electromagnetic free-field Lagrangians"—one advocated by Fermi, another by Schwinger, but all of which suffer from the formal defect just described.[24] One can temper the problem by adopting Procca theory (i.e., by pretending that "the photon has mass") for the purposes of calculation, and then "turning off $\varkappa$" at the end of the day.

One final remark, intended to clarify the "theoretical placement" of the preceding discussion, the distinction between what we have accomplished and what we have not: In physical applications of 3-dimensional vector analysis (whether to electrodynamics, to fluid dynamics, ...) one frequently gains very useful analytical leverage from *Helmholtz' theorem*, according to which every vector field $\boldsymbol{V}(x)$ can be expressed as the superposition

$$\boldsymbol{V}(x) = \boldsymbol{S}(x) + \boldsymbol{I}(x)$$

of a *solenoidal* field $\boldsymbol{S}(x)$ and an *irrotational* field $\boldsymbol{I}(x)$, where

$$\boldsymbol{\nabla}\cdot\boldsymbol{S} = 0 \Rightarrow \text{there exists an } \boldsymbol{\Omega} \text{ field such that } \boldsymbol{S} = \boldsymbol{\nabla}\times\boldsymbol{\Omega}$$
$$\boldsymbol{\nabla}\times\boldsymbol{I} = \boldsymbol{0} \Rightarrow \text{there exists an } \omega \text{ field such that } \boldsymbol{I} = \boldsymbol{\nabla}\omega$$

Helmholtz' theorem[25] speaks of a particular instance of the vastly more general *Hodge decomposition theorem*,[26] which—as it refers to vector fields on spacetime—asserts that every such field can be represented

$$V_\mu = S_\mu + I_\mu + V^0_\mu$$

where $\partial^\mu S_\mu = 0$, $\partial_\mu I_\nu - \partial_\nu I_\mu = 0$ and $\Box V^0_\mu = 0$, and that the representation is unique. In the preceding discussion we have been lead from these basic assumptions

$$c_1 + c_2 + c_3 = 0 \quad \text{and} \quad \varkappa \neq 0$$

[23] Change $U^\mu \to A^\mu$, $G^{\mu\nu} \to F^{\mu\nu}$.

[24] See also §4–9 in F. Rohrlich, *Classical Charged Particles* (1965).

[25] See R. B. McQuistan, *Scalar and Vector Fields: A Physical Interpretation* (1965) §11.5 for a detailed proof.

[26] See H. Flanders, *Differential Forms, with Applications to the Physical Sciences* (1963), p. 138.

to Procca fields U_μ from which the "irrotational" and "harmonic" terms are absent. We would expect such terms to be called into play if our assumptions were relaxed.[27]

Introduction to the "canonical formulation" of relativistic free-field theory. The following remarks are intended to place in useful perspective—and in that sense to be preparatory for—discussion of the Dirac equation. But they are of some independent interest, and will serve to indicate one of the portals through which algebra and group representation theory enter into field-theoretic discourse.

Let $m\ddot{x} = -U'(x)$ speak for the vast population of second-order differential equations which Nature dumps upon us in such variety. Through each initial point $x(0)$ pass continuously many solution curves $x(t)$, distinguished one from another by the values assigned to the initial velocity $\dot{x}(0)$. Write

$$\begin{aligned} \dot{x} &\equiv y \\ \dot{y} &= -\tfrac{1}{m}V'(x) \end{aligned}$$

The original *single* 2^{nd}*-order* equation has been displayed as a *pair of* 1^{st}*-order* equations; the advantages thus gained have to do with the circumstances that

- it is often easier to solve 1^{st}-order systems (even coupled systems of them);
- it is easier to comprehend the geometry of the solution space, since through each initial point $\big\{x(0), y(0)\big\}$ passes but a *single* solution $\big\{x(t), y(t)\big\}$.

It is just such a procedure (cunningly implemented) which sends

$$\text{Lagrangian formalism} \longrightarrow \text{Hamiltonian formalism}$$

Processes of the type

$$\begin{array}{l} G(\dddot{x}, \ddot{x}, \dot{x}, x) = 0 \\ \quad\downarrow \\ \ddot{x} = g(\ddot{x}, \dot{x}, x) \end{array} \quad \longrightarrow \quad \begin{cases} \dot{x} = y \\ \dot{y} = z \\ \dot{z} = g(z, y, x) \end{cases}$$

are called "reduction to canonical form," and are the frequently-encountered first step in work involving differential equations, whether ordinary or partial.

Look back again, in this light, to the Klein-Gordon equations

$$\big(\Box + \varkappa^2\big)\psi = 0 \quad \text{and its complex conjugate} \tag{20}$$

Define

$$\varkappa\,\psi_\mu \equiv \partial_\mu \psi \tag{33.1}$$

where the $\varkappa$-factor has been introduced in the presumption that $\varkappa \neq 0$, and in order to insure dimensional homogeneity: $[\psi_\mu] = [\psi]$. Then (20) becomes

$$\partial_\mu \psi^\mu + \varkappa\,\psi = 0 \tag{33.2}$$

[27] See, in this connection, the discussion on p. 91 in RELATIVISTIC CLASSICAL FIELDS (1973).

Equations (33) can be notated

$$\left.\begin{aligned} \varkappa\,\psi + \partial_\mu \psi^\mu &= 0 \\ \partial_\mu \psi - \varkappa\, g_{\mu\nu}\psi^\nu &= 0 \end{aligned}\right\} \tag{34}$$

or again

$$\begin{pmatrix} \varkappa & \partial_0 & \partial_1 & \partial_2 & \partial_3 \\ \partial_0 & -\varkappa & 0 & 0 & 0 \\ \partial_1 & 0 & +\varkappa & 0 & 0 \\ \partial_2 & 0 & 0 & +\varkappa & 0 \\ \partial_3 & 0 & 0 & 0 & +\varkappa \end{pmatrix} \begin{pmatrix} \psi \\ \psi^0 \\ \psi^1 \\ \psi^2 \\ \psi^3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \tag{35}$$

We began with a single partial differential equation of 2^{nd}-order, and ended up with a quintet of partial differential equations of 1^{st}-order.

Changing the sign of entries on the second (the 0^{th}) row, we find ourselves in position to write

$$\big(\boldsymbol{G}^\mu \partial_\mu + \varkappa \boldsymbol{I}\,\big)\Psi = 0 \tag{36}$$

where Ψ is the obvious 5-element column vector, and where the matrices $\boldsymbol{G}^\mu$ are defined

$$\left.\begin{aligned} \boldsymbol{G}^0 &\equiv \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{G}^1 \equiv \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \\ \boldsymbol{G}^2 &\equiv \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{G}^3 \equiv \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}\right\} \tag{37}$$

If we multiply i into (36) and define

$$\boldsymbol{\Gamma}^\mu \equiv i\,\boldsymbol{G}^\mu \tag{38}$$

then we obtain an equation

$$\big(\boldsymbol{\Gamma}^\mu \partial_\mu + i\varkappa \boldsymbol{I}\,\big)\Psi = 0 \tag{39}$$

which anticipates the structural design of the celebrated Dirac equation.

We will look in a moment to some of the important (and very beautiful) algebraic ramifications of the preceding formal development. But first we consider this question: Under what conditions can multi-component equations of the form

$$\big(\boldsymbol{\Gamma}^\mu \partial_\mu + i\varkappa \boldsymbol{I}\,\big)\Psi = 0 \quad : \quad \text{no special properties now ascribed to } \boldsymbol{\Gamma}^\mu$$

be brought within the compass of Lagrangian field theory? We expect the Lagrangian—if it exists—to be bilinear in starred and unstarred field variables, and look therefore to systems of the general type

$$\begin{aligned}\mathfrak{L} &= \Psi^{*a}_{\mu} P^{\mu}_{ab} \Psi^{b} + \Psi^{a} Q^{\mu}_{ab} \Psi^{b}_{\mu} + \varkappa \Psi^{a} G_{ab} \Psi^{b} \\ &= \Psi^{\dagger}_{\mu} \boldsymbol{P}^{\mu} \Psi + \Psi^{\dagger} \boldsymbol{Q}^{\mu} \Psi_{\mu} + \varkappa \Psi^{\dagger} \boldsymbol{G} \Psi\end{aligned}$$

where the dagger † signifies hermitian conjugation, where I have once again found it convenient to omit the commas which signify partial differentiation (writing Ψ_{μ} in place of $\Psi_{,\mu}$), and where the introduction of $\varkappa$ entails no loss of generality, but simplifies the endgame. The reality of $\mathfrak{L}$ entails that

$$\boldsymbol{G}^{\dagger} = \boldsymbol{G} \quad \text{and} \quad (\boldsymbol{G}^{\mu})^{\dagger} = \boldsymbol{Q}$$

Resolving $\boldsymbol{P}^{\mu}$ into its hermitian and antihermitian parts, we have

$$\begin{aligned}\boldsymbol{P}^{\mu} &= \boldsymbol{R}^{\mu} + i\tfrac{1}{2}\boldsymbol{S}^{\mu} \\ \boldsymbol{Q}^{\mu} &= \boldsymbol{R}^{\mu} - i\tfrac{1}{2}\boldsymbol{S}^{\mu}\end{aligned}$$

where $\boldsymbol{R}^{\mu}$ and $\boldsymbol{S}^{\mu}$ are both hermitian, and where the purpose of the $\frac{1}{2}$ will become clear almost immediately. The Lagrangian has now become

$$\mathfrak{L} = \underbrace{\big\{\Psi^{\dagger}_{\mu} \boldsymbol{R}^{\mu} \Psi + \Psi^{\dagger} \boldsymbol{R}^{\mu} \Psi_{\mu}\big\}}_{\text{gauge term: abandon}} + i\tfrac{1}{2}\big\{\Psi^{\dagger}_{\mu} \boldsymbol{S}^{\mu} \Psi - \Psi^{\dagger} \boldsymbol{S}^{\mu} \Psi_{\mu}\big\} + \varkappa \Psi^{\dagger} \boldsymbol{G} \Psi$$

The resulting field equations read

$$\boldsymbol{S}^{\mu} \Psi_{\mu} + i\,\boldsymbol{G} \Psi = 0 \quad \text{and hermitian conjugate}$$

We are brought to the conclusion that $\big(\boldsymbol{\Gamma}^{\mu} \partial_{\mu} + i\varkappa \boldsymbol{I}\big)\Psi = 0$ can be obtained from a Lagrangian if and only if there exists a non-singular hermitian matrix $\boldsymbol{G}$ such that $\boldsymbol{S}^{\mu} \equiv \boldsymbol{G}\boldsymbol{\Gamma}^{\mu}$ is hermitian. The Lagrangian can, in that case, be described

$$\mathfrak{L} = i\tfrac{1}{2}\big\{\Psi^{\dagger}_{\mu} \boldsymbol{G}\boldsymbol{\Gamma}^{\mu} \Psi - \Psi^{\dagger} \boldsymbol{G}\boldsymbol{\Gamma}^{\mu} \Psi_{\mu}\big\} + \varkappa \Psi^{\dagger} \boldsymbol{G} \Psi$$

and if we agree to write

$$\tilde{\Psi} \equiv \Psi^{\dagger} \boldsymbol{G} \tag{40}$$

becomes

$$\mathfrak{L} = i\tfrac{1}{2}\big\{\tilde{\Psi}_{\mu} \boldsymbol{\Gamma}^{\mu} \Psi - \tilde{\Psi} \boldsymbol{\Gamma}^{\mu} \Psi_{\mu}\big\} + \varkappa \tilde{\Psi} \Psi \tag{41}$$

At (39) we had

$$\boldsymbol{\Gamma}^{0} \equiv \begin{pmatrix} 0 & i & 0 & 0 & 0 \\ -i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{\Gamma}^{1} \equiv \begin{pmatrix} 0 & 0 & i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ +i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{\Gamma}^{2} \equiv \begin{pmatrix} 0 & 0 & 0 & i & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ +i & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad \boldsymbol{\Gamma}^{3} \equiv \begin{pmatrix} 0 & 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ +i & 0 & 0 & 0 & 0 \end{pmatrix}$$

and see now by inspection that a "hermitianizer" which works is

$$\boldsymbol{G} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} \tag{42}$$

I turn now to algebraic aspects of Klein-Gordon theory. Let (36) be notated

$$\boldsymbol{M}(\partial)\Psi = 0 \tag{43}$$

where

$$\boldsymbol{M}(\partial) \equiv \boldsymbol{G}^{\mu}\partial_{\mu} + \varkappa \boldsymbol{I} = \begin{pmatrix} \varkappa & \partial_0 & \partial_1 & \partial_2 & \partial_3 \\ -\partial_0 & \varkappa & 0 & 0 & 0 \\ \partial_1 & 0 & \varkappa & 0 & 0 \\ \partial_2 & 0 & 0 & \varkappa & 0 \\ \partial_3 & 0 & 0 & 0 & \varkappa \end{pmatrix} \tag{44}$$

and where my notation is intended to emphasize that $\boldsymbol{M}(\partial)$ has the character of a matrix-valued differential operator. Calculation gives

$$\det \boldsymbol{M}(\partial) = \varkappa^3\big(\,\square + \varkappa^2\,\big)$$

while matrix theory supplies the general proposition that

$$\boldsymbol{M}^{-1} = \frac{(\text{matrix of cofactors})^{\intercal}}{\det \boldsymbol{M}}$$

We conclude in the case at hand that

$$\boldsymbol{W}(\partial) \equiv \frac{(\text{matrix of cofactors})^{\intercal}}{\varkappa^3}$$

has the property that

$$\boldsymbol{W}(\partial) \cdot \big(\boldsymbol{G}^{\mu}\partial_{\mu} + \varkappa \boldsymbol{I}\,\big) = \big(\,\square + \varkappa^2\,\big)\boldsymbol{I} \tag{45}$$

From (45)—i.e., from the mere *existence* of such a $\boldsymbol{W}(\partial)$—it follows that if Ψ satisfies (43) then the components of Ψ *individually* satisfy the Klein-Gordon equation: $\big(\,\square + \varkappa^2\,\big)\psi^a = 0$.[28] In the context immediately at hand this is not

[28] Recall from electrodynamics the fairly tricky little argument which leads from the free-field equations

$$\partial_{\mu}F^{\mu\nu} = 0 \quad \text{and} \quad \partial^{\lambda}F^{\mu\nu} + \partial^{\mu}F^{\nu\lambda} + \partial^{\nu}F^{\lambda\mu} = 0$$

to the conclusion that

$$\square F^{\mu\nu} = 0 \quad : \quad \text{all } \mu \text{ and } \nu$$

startling news,[29] but it becomes powerfully informative when (as below) one undertakes to enlarge upon the present context.

One can proceed computationally from (45) to an explicit description of $\boldsymbol{W}(\partial)$—*Mathematica* 4.0, running on my PowerMac G3 took .0166667 seconds to do the job; by hand it takes a bit longer—but the result is so complicated, and (seemingly) so obscurely patterned, that it would serve no useful purpose to write it out; to do so would, however, make obvious this fact:

> The elements of $\boldsymbol{W}(\partial)$ present constants, ∂ and $\partial\partial$ operators, but no derivative operators of higher than second order. (46)

Let us now look *in general* to coupled first-order field equations of the design

$$\big(\boldsymbol{G}^\mu\partial_\mu + \varkappa\boldsymbol{I}\big)\Psi = 0 \qquad (36 \equiv 47)$$

—abandoning all special assumptions concerning

- the number of the field components ψ^a
- structural particulars of the matrices $\boldsymbol{G}^\mu$

but insisting that it remain possible to write (45); we look, in other words, to those instances of (47) for which it can be argued that if Ψ satisfies (47) then the individual components of Ψ satisfy the

$$\text{“Klein-Gordon condition”} \quad : \quad \big(\Box + \varkappa^2\big)\psi^a = 0 \quad (\text{all } a) \qquad (48)$$

To that end: write

$$\boldsymbol{W}(\partial) = \boldsymbol{A} + \boldsymbol{A}^\mu\partial_\mu + \boldsymbol{A}^{\mu\nu}\partial_\mu\partial_\nu + \cdots + \boldsymbol{A}^{\mu\nu\ldots\sigma}\partial_\mu\partial_\nu\cdots\partial_\sigma \qquad (49)$$

Then[30]

$$\begin{aligned}\boldsymbol{W}(\partial)\cdot\big(\boldsymbol{G}^\mu\partial_\mu + \varkappa\boldsymbol{I}\big) = \varkappa\boldsymbol{A} &+ \big(\varkappa\boldsymbol{A}^\mu + \boldsymbol{A}\boldsymbol{G}^\mu\big)\partial_\mu \\ &+ \big(2\varkappa\boldsymbol{A}^{\mu\nu} + [\boldsymbol{A}^\mu\boldsymbol{G}^\nu + \boldsymbol{A}^\nu\boldsymbol{G}^\mu]\big)\partial_\mu\partial_\nu \\ &+ \big(3\varkappa\boldsymbol{A}^{\mu\nu\rho} + [\boldsymbol{A}^{\mu\nu}\boldsymbol{G}^\rho + \boldsymbol{A}^{\nu\rho}\boldsymbol{G}^\mu + \boldsymbol{A}^{\rho\mu}\boldsymbol{G}^\nu]\big)\partial_\mu\partial_\nu\partial_\rho \\ &+ \cdots \\ = \varkappa^2\boldsymbol{I} + 2g^{\mu\nu}\boldsymbol{I}\partial_\mu\partial_\nu \quad &: \quad \text{REQUIRED}\end{aligned}$$

[29] Recall from (35) how the ψ^a were defined, and observe that if ψ satisfies the K-G equation then certainly the functions $\partial^\mu\psi$ do.

[30] We intend term-wise comparison of the coefficients of $\partial_\mu\partial_\nu\cdots\partial_\sigma$, and therefore (to avoid implications of the circumstance that such expressions, by their total symmetry, wear multiple aliases) understand it to be the case that

$$\sum_{\text{Greek indices}} \text{ are subject to the constraint that } \mu \leqslant \nu \leqslant \cdots \leqslant \sigma$$

Additionally, we recognize the symmetry of $g^{\mu\nu}$ and the total symmetry of $\boldsymbol{A}^{\mu\nu\ldots\sigma}$. This accounts for the intrusive integers and some otherwise inexplicable features of the following equations.

gives

$$\begin{aligned}
\varkappa \boldsymbol{A} &= \varkappa^2 \boldsymbol{I} \\
\varkappa \boldsymbol{A}^\mu + \boldsymbol{A}\boldsymbol{G}^\mu &= \boldsymbol{0} \\
2\varkappa \boldsymbol{A}^{\mu\nu} + [\boldsymbol{A}^\mu \boldsymbol{G}^\nu + \boldsymbol{A}^\nu \boldsymbol{G}^\mu] &= 2g^{\mu\nu}\boldsymbol{I} \\
3\varkappa \boldsymbol{A}^{\mu\nu\rho} + [\boldsymbol{A}^{\mu\nu}\boldsymbol{G}^\rho + \boldsymbol{A}^{\nu\rho}\boldsymbol{G}^\mu + \boldsymbol{A}^{\rho\mu}\boldsymbol{G}^\nu] &= \boldsymbol{0} \\
&\vdots
\end{aligned}$$

which can be solved serially, to give

$$\boldsymbol{A} = \varkappa \boldsymbol{I} \tag{50.0}$$

$$\boldsymbol{A}^\mu = -\boldsymbol{G}^\mu \tag{50.1}$$

$$\boldsymbol{A}^{\mu\nu} = \tfrac{1}{\varkappa}\big\{g^{\mu\nu}\boldsymbol{I} + \tfrac{1}{2}\big[\boldsymbol{G}^\mu\boldsymbol{G}^\nu + \boldsymbol{G}^\nu\boldsymbol{G}^\mu\big]\big\} \tag{50.2}$$

$$\boldsymbol{A}^{\mu\nu\rho} = -\tfrac{1}{3\varkappa^2} \sum_{\text{cyclic permutations}} \big\{g^{\mu\nu}\boldsymbol{G}^\rho + \tfrac{1}{2}\big[\boldsymbol{G}^\mu\boldsymbol{G}^\nu + \boldsymbol{G}^\nu\boldsymbol{G}^\mu\big]\boldsymbol{G}^\rho\big\} \tag{50.3}$$

$$\vdots$$

It is clear from the argument that gave (50) that if all $\boldsymbol{A}$-matrices of order m vanish, then so also do all $\boldsymbol{A}$-matrices of order $n > m$. The series (49) then truncates at order $m-1$.

Suppose, for example, that the matrices $\boldsymbol{G}^\mu$ have the property that

$$\sum_{\text{cyclic permutations}} \big\{g^{\mu\nu}\boldsymbol{G}^\rho + \tfrac{1}{2}\big[\boldsymbol{G}^\mu\boldsymbol{G}^\nu + \boldsymbol{G}^\nu\boldsymbol{G}^\mu\big]\boldsymbol{G}^\rho\big\} = \boldsymbol{0} \quad : \quad \text{all } \mu,\, \nu,\, \rho \tag{51}$$

Then introduction of (50) into (49) gives

$$\boldsymbol{W}(\partial) = \tfrac{1}{\varkappa}\big(\Box + \varkappa^2\big)\boldsymbol{I} - \boldsymbol{G}^\mu\partial_\mu + \tfrac{1}{2\varkappa}\big[\boldsymbol{G}^\mu\boldsymbol{G}^\nu + \boldsymbol{G}^\nu\boldsymbol{G}^\mu\big]\partial_\mu\partial_\nu \tag{52}$$

One can with labor show that the 5×5 matrices $\boldsymbol{G}^\mu$ introduced at (37) do in fact satisfy (51). And that (52) provides a structured account of precisely the "obscurely patterned" matrix calculated by *Mathematica*!

Conditions (51) can be shown[31] to be expressible

$$\boldsymbol{\Gamma}^\mu\boldsymbol{\Gamma}^\nu\boldsymbol{\Gamma}^\rho + \boldsymbol{\Gamma}^\rho\boldsymbol{\Gamma}^\nu\boldsymbol{\Gamma}^\mu = g^{\mu\nu}\boldsymbol{\Gamma}^\rho + g^{\rho\nu}\boldsymbol{\Gamma}^\mu$$

where the $\boldsymbol{\Gamma}$-matrices are those introduced at (38). The latter relations give rise to what is called "Kemmer-Duffin algebra," which was first studied in the late 1930's.[32] In higher order one encounters algebraic structures of increasing complexity, about which progressively less is known.

---

[31] See pp. 133–136 in RELATIVISTIC CLASSICAL FIELDS (1973).

[32] R. J. Duffin, "On the characteristic matrices of covariant systems," Phys. Rev. **54**, 1114 (1938); N. Kemmer, "The particle aspect of meson theory," Proc. Roy. Soc. (London) **173A**, 91 (1939). The topic had been explored already by G. Petiau in 1936.

Simplest case: the Dirac equation. The theory of a scalar Klein-Gordon field, when expressed in canonical form, was seen to lead to a 5-component Ψ-field, and to an instance of (49) which truncates in second order. More complicated covariant field theories result when we demand truncation in third, fourth or higher order. But what happens if we insist upon truncation in second (the lowest possible) order?

To do so is to impose upon the $\boldsymbol{G}$-matrices the requirement that

$$g^{\mu\nu}\boldsymbol{I} + \tfrac{1}{2}\big[\boldsymbol{G}^\mu\boldsymbol{G}^\nu + \boldsymbol{G}^\nu\boldsymbol{G}^\mu\big] = \boldsymbol{0} \quad : \quad \text{all } \mu \text{ and } \nu$$

which in terms of the $\boldsymbol{\Gamma}$-matrices are those introduced at (38) becomes[33]

$$\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu + \boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\mu = 2g^{\mu\nu}\boldsymbol{I} \quad : \quad \text{all } \mu \text{ and } \nu \tag{53.1}$$

The generic field equation (36) reads[34]

$$\big(\boldsymbol{G}^\mu\partial_\mu + \varkappa\boldsymbol{I}\big)\psi = 0$$

and introduction of (50) into (49) now gives $\boldsymbol{W}(\partial) = \varkappa\boldsymbol{I} - \boldsymbol{G}^\mu\partial_\mu$ whence

$$\big(-\boldsymbol{G}^\mu\partial_\mu + \varkappa\boldsymbol{I}\big)\big(\boldsymbol{G}^\nu\partial_\nu + \varkappa\boldsymbol{I}\big)\psi = \big(\Box + \varkappa^2\big)\psi$$

In $\boldsymbol{\gamma}$-notation the preceding equations read

$$\big(\boldsymbol{\gamma}^\mu\partial_\mu + i\varkappa\boldsymbol{I}\big)\psi = 0 \tag{53.2}$$

$$\big(\boldsymbol{\gamma}^\mu\partial_\mu - i\varkappa\boldsymbol{I}\big)\big(\boldsymbol{\gamma}^\nu\partial_\nu + i\varkappa\boldsymbol{I}\big)\psi = \big(\Box + \varkappa^2\big)\psi \tag{53.3}$$

Equation (53.2) is precisely the *Dirac equation*, put forward in (1927) by P. A. M. Dirac on the basis of quite a different set of considerations.[35] Dirac was able and content simply to pluck out of thin air a quartet of complex 4×4

[33] I will, in fact, write $\boldsymbol{\gamma}^\mu$ in place of $\boldsymbol{\Gamma}^\mu$—partly to underscore the fact that we have particularized a generic situation, but mainly to come into agreement with long-established notational convention.

[34] Since ψ is, in the generic case, not a preempted symbol (no scalar field is a presumed player) I will henceforth write ψ where formerly we wrote Ψ; the latter symbol will see service again when we turn to the *transformational* aspects of the theory.

[35] "The quantum theory of the electron," Proc. Roy. Soc. (London) **117A**, 610 (1928) and **118A**, 351 (1928); Chapter 11, *The Principles of Quantum Mechanics* (4^{th} edition 1958). Dirac noted that, while for a non-relativistic free particle $E = \frac{1}{2m}(p_1^2 + p_2^2 + p_3^2)$ goes over by Schrödinger quantization into a perfectly acceptable Schrödinger equation, in relativity one has

$$g^{\mu\nu}p_\mu p_\nu = (E/c)^2 - (p_1^2 + p_2^2 + p_3^2) = (mc)^2$$

which by the same procedure yields a quantum mechanically unacceptable second time derivative. Writing $E = \pm\sqrt{(mc)^2 + p_1^2 + p_2^2 + p_3^2}$ does not help,

matrices which satisfy (53.1), but soon it was shown (by Pauli) that

- the complex matrices of *least* dimension which satisfy (53.1) are 4×4;
- if 4×4 quartets $\tilde{\boldsymbol{\gamma}}^\mu$ and $\boldsymbol{\gamma}^\mu$ both satisfy (53.1) then necessarily there exists a $\boldsymbol{S}$ such that $\tilde{\boldsymbol{\gamma}}^\mu = \boldsymbol{S}\boldsymbol{\gamma}^\mu\boldsymbol{S}^{-1}$.

The latter fact ensures that we do not have "disjoint realizations" to contend with, and accounts for the fact that one encounters diverse realizations in the literature. The following "Bjorken & Drell" realization

$$\left.\begin{aligned} \boldsymbol{\gamma}^0 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \boldsymbol{\gamma}^1 = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\ \boldsymbol{\gamma}^2 &= \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{\gamma}^3 = \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix} \end{aligned}\right\} \tag{54}$$

can in terms of the Pauli matrices

$$\boldsymbol{I} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \boldsymbol{\sigma}_1 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_2 \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \boldsymbol{\sigma}_3 \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

be notated

$$\boldsymbol{\gamma}^0 = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix}, \qquad \boldsymbol{\gamma}^1 = \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_1 & \boldsymbol{0} \end{pmatrix}$$

$$\boldsymbol{\gamma}^2 = \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{\sigma}_2 \\ \boldsymbol{\sigma}_2 & \boldsymbol{0} \end{pmatrix}, \qquad \boldsymbol{\gamma}^3 = \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{\sigma}_3 \\ \boldsymbol{\sigma}_3 & \boldsymbol{0} \end{pmatrix}$$

and enjoy fairly wide usage.[36] Note that while the Pauli matrices are hermitian, the matrices $\boldsymbol{\gamma}^m$ $(m = 1, 2, 3)$ are antihermitian. They can, however, be

---

for to do so introduces a sign ambiguity, and fractures the energy-momentum symmetry upon which relativity insists. Dirac's idea was to achieve a symmetry-preserving first-order expression by *factoring* $g^{\mu\nu}p_\mu p_\nu$, writing

$$g^{\mu\nu}p_\mu p_\nu = (\gamma^\mu p_\mu)(\gamma^\nu p_\nu)$$

even though this entails that the γ^μ be "hypernumbers" (matrices) constrained to satisfy

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$$

Dirac appears to have been unaware that "formal factorization of quadratic forms" is a topic which had been explored by W. K. Clifford (1845–1879) already in 1876, and that γ-algebra ("Dirac algebra") provides a special instance of a *Clifford algebra*.

[36] They are, in particular, employed by David Griffiths; see p. 216 of his *Introduction to Elementary Particles* (1987).

"hermitianized," which is to say: there exists a non-singular hermitian $\boldsymbol{G}$ with the property that $\boldsymbol{G}\boldsymbol{\gamma}^\mu$ is hermitian (all μ). The $\boldsymbol{G}$ which transparently does the trick is

$$\boldsymbol{G} \equiv \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix} \text{ which, as it happens, is just } \boldsymbol{\gamma}^0 \tag{55}$$

We are in position now to state (see again (40/41)) that *the Dirac theory admits of Lagrangian formulation*, and that the Lagrangian in question can be written

$$\begin{aligned} \mathfrak{L} &= -\hbar c\Big\{ i\tfrac{1}{2}\big\{\tilde{\psi}_\mu \boldsymbol{\gamma}^\mu \psi - \tilde{\psi}\boldsymbol{\gamma}^\mu \psi_\mu\big\} + \varkappa\, \tilde{\psi}\psi \Big\} \\ &= \hbar c\Big\{ \frac{\tilde{\psi}_\mu \boldsymbol{\gamma}^\mu \psi - \tilde{\psi}\boldsymbol{\gamma}^\mu \psi_\mu}{2i} - \varkappa\, \tilde{\psi}\psi \Big\} \end{aligned} \tag{56}$$

with $\tilde{\psi} \equiv \psi^\dagger \boldsymbol{G}$. The $\hbar c$-factor has been introduced in order to ensure that $[\mathfrak{L}] = \text{energy}/(\text{length})^3$, and in the presumption that $[\tilde{\psi}\psi] = 1/(\text{length})^4$, and the minus sign is physically inconsequential/cosmetic.

Drawing upon (1–41) we find that the energy density implicit in (56) might be described

$$\begin{aligned} \mathcal{E} &= \tilde{\psi}_0 \frac{\partial \mathfrak{L}}{\partial \tilde{\psi}_0} + \frac{\partial \mathfrak{L}}{\partial \psi_0}\psi_0 - \mathfrak{L} \\ &= \hbar c\Big\{ \frac{\tilde{\psi}_0 \boldsymbol{\gamma}^0 \psi - \tilde{\psi}\boldsymbol{\gamma}^0 \psi_0}{2i} \Big\} - \mathfrak{L} \end{aligned} \tag{57}$$

but $\mathfrak{L}$ ← can be dropped: see below.

The generic canonical Lagrangian (41) yields field equations

$$i\boldsymbol{\Gamma}^\mu \Psi_\mu = +\varkappa\Psi \quad \text{and} \quad i\tilde{\Psi}_\mu \boldsymbol{\Gamma}^\mu = -\varkappa\tilde{\Psi}$$

which when inserted back into (41) give

$$\begin{aligned} \mathfrak{L} &= \tfrac{1}{2}\big\{ -\tilde{\Psi}\varkappa\Psi - \tilde{\Psi}\varkappa\Psi \big\} + \varkappa\tilde{\Psi}\Psi \\ &= 0 \quad \text{in numerical value} \end{aligned} \tag{58}$$

This result is not as strange as it might at first appear, for it is the *functional form*—not the numerical value—of $\mathfrak{L}$ which is of dynamical consequence. Besides, such a result is familiar already from mechanics: Hamilton's canonical equations

$$\begin{aligned} \dot{q} &= +\tfrac{\partial}{\partial p} H \\ \dot{p} &= -\tfrac{\partial}{\partial q} H \end{aligned}$$

can be obtained as "Lagrange equations" from the "meta-Lagrangian"

$$L(\dot{q}, \dot{p}, q, p) \equiv \tfrac{1}{2}\big\{\dot{q}p - q\dot{p}\big\} - H(p, q)$$

and when inserted back *into* the meta-Lagrangian give

$$L = \tfrac{1}{2}\{q\tfrac{\partial}{\partial q} + p\tfrac{\partial}{\partial p}\}H - H$$

But if H is homogeneous of degree two then $\{q\frac{\partial}{\partial q} + p\frac{\partial}{\partial p}\}H = 2H$ (by Euler's theorem), and we have

$$L = 0 \quad \text{in numerical value}$$

And homogeneity of degree two ($H = ap^2 + bpq + cq^2$) is precisely the condition required to ensure *linearity* of the canonical equations. Notice that because $\dot{q}$ and $\dot{p}$ enter linearly into the design of L, it is not possible to construct an associated "meta-Hamiltonian." Similar remarks pertain whenever derivatives enter linearly into the design of a Lagrangian (or Lagrange density)... as in canonical theories they invariably and characteristically do.

Noether's theorem supplies this satisfying account of the energy of the meta-Lagrangian system just mentioned:

$$\begin{aligned} E &= \{\dot{q}\tfrac{\partial}{\partial \dot{q}} + \dot{p}\tfrac{\partial}{\partial \dot{p}}\}L - L \\ &= \tfrac{1}{2}\{\dot{q}p - \dot{p}q\} - \Big\{\tfrac{1}{2}\{\dot{q}p - q\dot{p}\} - H(p,q)\Big\} \\ &= H(p,q) \end{aligned}$$

The (conserved) total energy present in a free Dirac field can be described

$$\begin{aligned} E &= \iiint \mathcal{E}\, dx^1 dx^2 dx^3 \\ &= \tfrac{1}{2}\hbar c \iiint i\Big\{\tilde{\psi}\boldsymbol{\gamma}^0\psi_0 - \tilde{\psi}_0\boldsymbol{\gamma}^0\psi\Big\}\, dx^1 dx^2 dx^3 \end{aligned}$$

But the field equations supply

$$\begin{aligned} i\boldsymbol{\gamma}^0\psi_0 &= -i\boldsymbol{\gamma}\cdot\overrightarrow{\boldsymbol{\nabla}}\psi + \varkappa\,\psi \\ i\tilde{\psi}_0\boldsymbol{\gamma}^0 &= -i\tilde{\psi}\boldsymbol{\gamma}\cdot\overleftarrow{\boldsymbol{\nabla}} - \varkappa\,\psi \end{aligned}$$

so

$$\begin{aligned} E &= \hbar c \iiint \Big\{ - i\tfrac{1}{2}\tilde{\psi}\big[\boldsymbol{\gamma}\cdot\overrightarrow{\boldsymbol{\nabla}} - \boldsymbol{\gamma}\cdot\overleftarrow{\boldsymbol{\nabla}}\big]\psi + \varkappa\,\tilde{\psi}\psi\Big\}\, dx^1 dx^2 dx^3 \\ &= \hbar c \iiint \tilde{\psi}\Big\{ - i\boldsymbol{\gamma}\cdot\overrightarrow{\boldsymbol{\nabla}} + \varkappa\Big\}\psi\, dx^1 dx^2 dx^3 + i\tfrac{1}{2}\hbar c \iiint \overrightarrow{\boldsymbol{\nabla}}\cdot(\tilde{\psi}\boldsymbol{\gamma}\,\psi)\, dx^1 dx^2 dx^3 \end{aligned}$$

Integration by parts has yielded a second term which can (by the divergence theorem) be expressed $\iint(\tilde{\psi}\boldsymbol{\gamma}\,\psi)\cdot d\boldsymbol{\sigma}$ and abandoned. The resulting expression

$$E = \hbar c \iiint \tilde{\psi}\Big\{ - i\boldsymbol{\gamma}\cdot\boldsymbol{\nabla} + \varkappa\Big\}\psi\, dx^1 dx^2 dx^3 \tag{59}$$

is assigned the role of a "Hamiltonian" in some approaches[37] to *quantization* of the classical Dirac field.

The argument which gave (59) is not special to the Dirac field; it pertains to *all* canonically formulated Lagrangian field theories. It pertains in particular to the 5-component canonical formulation of the theory of a scalar Klein-Gordon field, and could be used to cast interesting new light upon (16.1). I shall, however, not pursue the detailed implications of this remark.

Dirac's theory, when looked upon as relativistic quantum mechanics of a particle ("quantum theory of the electron" in Dirac's phrase) proved to be a theory the explanatory power of which is surpassed only by its elegance. But my subject is classical field theory, so I must be content to leave all of that to Dirac's Chapters 11 & 12—to the advanced quantum textbooks.

Lorentz transform properties of multi-component fields. Revert again to the generic case (41):

$$\mathfrak{L} = i\tfrac{1}{2}\{\tilde{\psi}_\mu \boldsymbol{\Gamma}^\mu \psi - \tilde{\psi}\boldsymbol{\Gamma}^\mu \psi_\mu\} + \varkappa \tilde{\psi}\,\psi$$

with $\tilde{\psi} \equiv \psi^\dagger \boldsymbol{G}$. More explicitly (and for present purposes more conveniently)

$$\mathfrak{L} = i\tfrac{1}{2}\{\psi^\dagger_\mu \boldsymbol{G}\boldsymbol{\Gamma}^\mu \psi - \psi^\dagger \boldsymbol{G}\boldsymbol{\Gamma}^\mu \psi_\mu\} + \varkappa \psi^\dagger \boldsymbol{G}\psi \tag{60}$$

Lorentz covariance of the resulting field equations requires that each of the terms which enter into the construction of $\mathfrak{L}$ be Lorentz invariant. Enlarging upon our experience with tensor fields, we postulate it to be the case that

$$\Lambda: \quad x \to X = \boldsymbol{\Lambda} x \tag{61.0}$$

causes the components of ψ to fold linearly among themselves:

$$\Lambda: \quad \psi^a(x) \to \Psi^a(X) = U^a{}_b(\Lambda)\psi^b(x(X)) \tag{61.1}$$

which we abbreviate

$$\Lambda: \quad \psi \to \Psi = \boldsymbol{U}(\Lambda)\psi$$

Specific instances of such field-transformation laws have been encountered already at (5/6/7). It follows from (61.1) that first partials of the multi-component field necessarily transform

$$\begin{aligned}\Lambda: \quad \psi^a_{,\mu}(x) \to \Psi^a_{,\mu}(X) &= U^a{}_b(\Lambda)\tfrac{\partial x^\nu}{\partial X^\mu}\psi^b_{,\nu}(x(X)) \\ &= U^a{}_b(\Lambda)\,\Lambda^\nu{}_\mu \psi^b_{,\nu}(x(X))\end{aligned} \tag{61.2}$$

which we abbreviate

$$\Lambda: \quad \psi_\mu \to \Psi_\mu = \boldsymbol{U}(\Lambda)\,\Lambda^\nu{}_\mu \psi_\nu$$

[37] Recall (1–81) and see Schweber,[7] p. 220.

Looking back now to (60): we have

$$\begin{aligned}\mathfrak{L} &= i\tfrac{1}{2}\big\{\Psi^\dagger_\mu \boldsymbol{G\Gamma}^\mu\Psi - \Psi^\dagger \boldsymbol{G\Gamma}^\mu\Psi_\mu\big\} + \varkappa\,\Psi^\dagger\boldsymbol{G}\Psi \\ &= i\tfrac{1}{2}\big\{\psi^\dagger_\nu \Lambda^\nu{}_\mu \boldsymbol{U}^\dagger\boldsymbol{G\Gamma}^\mu\boldsymbol{U}\psi - \psi^\dagger\boldsymbol{U}^\dagger\boldsymbol{G\Gamma}^\mu\boldsymbol{U}\,\Lambda^\nu{}_\mu\psi_\nu\big\} + \varkappa\,\psi^\dagger\boldsymbol{U}^\dagger\boldsymbol{GU}\psi\end{aligned}$$

and insist upon

$$= i\tfrac{1}{2}\big\{\psi^\dagger_\nu \boldsymbol{G\Gamma}^\nu\psi - \psi^\dagger\boldsymbol{G\Gamma}^\nu\psi_\nu\big\} + \varkappa\,\psi^\dagger\boldsymbol{G}\psi$$

We are led thus to the requirements

$$\begin{aligned}\boldsymbol{U}^\dagger\boldsymbol{GU} &= \boldsymbol{G} \\ \Lambda^\nu{}_\mu\boldsymbol{U}^\dagger\boldsymbol{G\Gamma}^\mu\boldsymbol{U} &= \boldsymbol{G\Gamma}^\nu\end{aligned}$$

The former requirement can be written

$$\boldsymbol{U}^{-1} = \boldsymbol{G}^{-1}\boldsymbol{U}^\dagger\boldsymbol{G} \tag{62.1}$$

which is reminiscent of the Lorentz condition $\boldsymbol{\Lambda}^{-1} = \boldsymbol{g}^{-1}\boldsymbol{\Lambda}^{\mathsf{T}}\boldsymbol{g}$ encountered at (3), and can be used to bring the second set of required relations to the form

$$\boldsymbol{U}^{-1}\boldsymbol{\Gamma}^\mu\boldsymbol{U} = \Lambda^\mu{}_\nu\boldsymbol{\Gamma}^\nu \tag{62.2}$$

How do (62) check out as they relate to the canonical formulation of a scalar Klein-Gordon field φ? In that case

$$\psi = \begin{pmatrix}\psi^1\\ \psi^2\\ \psi^3\\ \psi^4\\ \psi^5\end{pmatrix} = \begin{pmatrix}\varphi\\ \partial^0\varphi\\ \partial^1\varphi\\ \partial^2\varphi\\ \partial^3\varphi\end{pmatrix} \tag{63}$$

The top component transforms as a scalar, the bottom four as components of a contravariant 4-vector… so we have

$$\boldsymbol{U}(\Lambda) = \begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & \begin{pmatrix}\Lambda^0{}_0 & \Lambda^0{}_1 & \Lambda^0{}_2 & \Lambda^0{}_3\\ \Lambda^1{}_0 & \Lambda^1{}_1 & \Lambda^1{}_2 & \Lambda^1{}_3\\ \Lambda^2{}_0 & \Lambda^2{}_1 & \Lambda^2{}_2 & \Lambda^2{}_3\\ \Lambda^3{}_0 & \Lambda^3{}_1 & \Lambda^3{}_2 & \Lambda^3{}_3\end{pmatrix}\end{pmatrix} \tag{64}$$

The matrix is in this case real so $\boldsymbol{U}^\dagger = \boldsymbol{U}^{\mathsf{T}}$, and (62.1) can be notated

$$\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Lambda}^{\mathsf{T}}\end{pmatrix}\boldsymbol{G}\begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{\Lambda}\end{pmatrix} = \boldsymbol{G} \tag{65}$$

But we found at (42) that the hermitianizer can in that case be described

$$\boldsymbol{G} = \begin{pmatrix}1 & \mathbf{0}\\ \mathbf{0} & \boldsymbol{g}\end{pmatrix} \tag{66}$$

so (62.1) is an immediate consequence of (3). To establish (62.2)—and at the same time to avoid the complexities latent in $\boldsymbol{U}^{-1}$—it is sufficient to show that for all 4-vectors a_μ

$$a_\mu \boldsymbol{\Gamma}^\mu \boldsymbol{U} = a_\mu \Lambda^\mu{}_\nu \boldsymbol{U} \boldsymbol{\Gamma}^\nu \tag{67}$$

which is an inelegantly straightforward computational assignment; recalling from (37/38) the definitions of the $\boldsymbol{\Gamma}$-matrices, we look to the left side of (67) and compute

$$\begin{pmatrix} 0 & a_0 & a_1 & a_2 & a_3 \\ -a_0 & 0 & 0 & 0 & 0 \\ +a_1 & 0 & 0 & 0 & 0 \\ +a_2 & 0 & 0 & 0 & 0 \\ +a_3 & 0 & 0 & 0 & 0 \end{pmatrix} \boldsymbol{U} = \begin{pmatrix} 0 & a_\mu \Lambda^\mu{}_0 & a_\mu \Lambda^\mu{}_1 & a_\mu \Lambda^\mu{}_2 & a_\mu \Lambda^\mu{}_3 \\ -a^0 & 0 & 0 & 0 & 0 \\ -a^1 & 0 & 0 & 0 & 0 \\ -a^2 & 0 & 0 & 0 & 0 \\ -a^3 & 0 & 0 & 0 & 0 \end{pmatrix}$$

while the expression on the right side of (67) yields

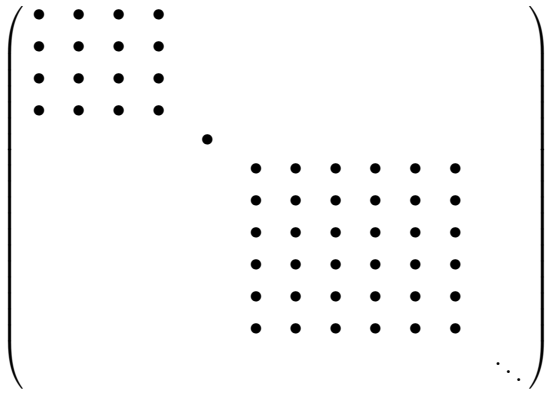
$$a_\mu \Lambda^\mu{}_0 \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ -\Lambda^0{}_0 & 0 & 0 & 0 & 0 \\ -\Lambda^1{}_0 & 0 & 0 & 0 & 0 \\ -\Lambda^2{}_0 & 0 & 0 & 0 & 0 \\ -\Lambda^3{}_0 & 0 & 0 & 0 & 0 \end{pmatrix} + a_\mu \Lambda^\mu{}_1 \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ +\Lambda^0{}_1 & 0 & 0 & 0 & 0 \\ +\Lambda^1{}_1 & 0 & 0 & 0 & 0 \\ +\Lambda^2{}_1 & 0 & 0 & 0 & 0 \\ +\Lambda^3{}_1 & 0 & 0 & 0 & 0 \end{pmatrix} + \cdots$$

The top row has precisely the desired structure, and all the 0's are correctly placed. Looking finally to the elements of the first column, the design of the Lorentz metric entails

$$\begin{aligned} & - a_\mu \Lambda^\mu{}_0 \Lambda^\nu{}_0 + a_\mu \Lambda^\mu{}_1 \Lambda^\nu{}_1 + a_\mu \Lambda^\mu{}_2 \Lambda^\nu{}_2 + a_\mu \Lambda^\mu{}_3 \Lambda^\nu{}_3 \\ = & - a_\mu \Lambda^{\mu 0} \Lambda^\nu{}_0 - a_\mu \Lambda^{\mu 1} \Lambda^\nu{}_1 - a_\mu \Lambda^{\mu 2} \Lambda^\nu{}_2 - a_\mu \Lambda^{\mu 3} \Lambda^\nu{}_3 \\ = & - a_\mu g^{\mu\nu} \\ = & - a^\nu \end{aligned}$$

which completes the demonstration.

Matrices $\boldsymbol{U}(\Lambda)$ which have (or by similarity transformation can be made to have) the block design



(as exemplified by (64)) provide "reducible" representations of the group in question—here: the Lorentz group. Matrices which do not admit of such reduction (such as the sub-matrices found on the diagonal) are said to be "irreducible." It is with the enumeration/description of the irreducible representations that group representation theory is principally concerned.[38]

We turn now from Klein-Gordon theory to the more interesting problems presented by the Dirac theory. The Dirac field

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix}$$

came to be a 4-component complex field because (it was reported without proof) the least-dimensional representation of

$$\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu + \boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\mu = 2g^{\mu\nu}\boldsymbol{I} \quad : \quad \text{all } \mu \text{ and } \nu \tag{53.1}$$

is complex 4×4. It is tempting to suppose that the components of ψ transform

$$\psi \to \Psi = \boldsymbol{U}(\Lambda)\psi \tag{68}$$

as components of a 4-vector, but this cannot be the case: $\boldsymbol{U}$ was constrained at (62.1) to satisfy

$$\boldsymbol{U}^{-1} = \boldsymbol{G}^{-1}\boldsymbol{U}^\dagger\boldsymbol{G}$$

and (see again (55))

$$\boldsymbol{G} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{is non-Lorentzian } (\neq \boldsymbol{g})$$

Evidently the Dirac field ψ transforms, according to (68), as a new kind of object—a "4-spinor," and the "4-ness" of the matter has nothing (or at least nothing obvious) to do with the 4-dimensionality of spacetime.

Complex representations of the Lorentz group are recommended to our attention by the simplest of considerations: set up the association

$$x = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \quad \longleftrightarrow \quad \boldsymbol{x} = \begin{pmatrix} x^0 + x^3 & x^1 - ix^2 \\ x^1 + ix^2 & x^0 - x^3 \end{pmatrix} \; : \; \text{hermitian}$$

[38] See (for example) F. D. Murnaghan, *The Theory of Group Representations* (1938), Chapter 2.

and notice that

$$\det \boldsymbol{x} = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2 = g_{\mu\nu}x^\mu x^\nu = (x,x)$$

is invariant under

$$\boldsymbol{x} \to \boldsymbol{X} = \boldsymbol{U}^{-1}\boldsymbol{x}\boldsymbol{U} \ : \ \boldsymbol{U} \text{ is complex } 2\times 2$$

The complex matrix $\boldsymbol{U}$ contains eight assignable constants, but if we impose the "unimodularity" condition $\det \boldsymbol{U} = 1 + i0$ that number is reduced to six. We may expect to set up an association between elements $\boldsymbol{\Lambda}$ of the 6-parameter Lorentz group and elements $\boldsymbol{S}$ of the 6-parameter group $SU(2)$ of unimodular complex 2×2 matrices (all of which are, as it happens, automatically unitary):

$$\boldsymbol{\Lambda} \quad \longleftrightarrow \quad \boldsymbol{S}(\Lambda) \ : \ \text{element of } SU(2)$$

Then

$$\boldsymbol{x} \to \boldsymbol{X} = \boldsymbol{S}^{-1}\boldsymbol{x}\boldsymbol{S} \quad \text{provides a representation of} \quad x \to X = \boldsymbol{\Lambda}x \tag{69.1}$$

Two-component complex objects $\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix}$ which participate by transforming

$$\psi \to \Psi = \boldsymbol{S}\,\psi \tag{69.2}$$

are said to transform as "2-spinors." In it important to notice that the association $\boldsymbol{\Lambda} \leftrightarrow \boldsymbol{S}(\Lambda)$ cannot be unique (is "biunique"), for $\boldsymbol{S}(\Lambda)$ enters quadratically into the equation which describes $\boldsymbol{x} \to \boldsymbol{X}$: the matrices $\pm\boldsymbol{S}$ both achieve the same effect, though the same cannot be said of (69.2).

At (68) we are presented with occasion to develop a grander variant of that same general kind of mathematics—a variant in which the design of $\boldsymbol{U}(\Lambda)$ is controlled by the statements

$$\boldsymbol{U}^{-1} = \boldsymbol{G}^{-1}\boldsymbol{U}^\dagger\boldsymbol{G} \tag{70.1}$$

$$\boldsymbol{U}^{-1}\boldsymbol{\gamma}^\mu\boldsymbol{U} = \Lambda^\mu{}_\nu\boldsymbol{\gamma}^\nu \tag{70.2}$$

Both equations are invariant under $\boldsymbol{U} \to -\boldsymbol{U}$, so we expect the association $\boldsymbol{\Lambda} \leftrightarrow \boldsymbol{U}(\Lambda)$ to be once again biunique. Note also that the metric structure $\boldsymbol{g}$ of spacetime enters (70.2) once through the relations $\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu + \boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\mu = 2g^{\mu\nu}\boldsymbol{I}$ that control the design of the $\boldsymbol{\gamma}$-matrices, and once through the equation $\boldsymbol{\Lambda}^{-1} = \boldsymbol{g}^{-1}\boldsymbol{\Lambda}^\intercal\boldsymbol{g}$ that controls the design of the numbers $\Lambda^\mu{}_\nu$. If the $\boldsymbol{G}$-factors were absent from (70.1) then that equation would assert the *unitarity* of $\boldsymbol{U}$; noting that if we define

$$(\varphi,\psi) \equiv \varphi^\dagger\boldsymbol{G}\,\psi = \tilde{\varphi}\,\psi$$

and write $\psi \to \Psi = \boldsymbol{U}\psi$ then (70.1), written $\boldsymbol{G} = \boldsymbol{U}^\dagger\boldsymbol{G}\boldsymbol{U}$, entails

$$(\varphi,\psi) = (\Phi,\Psi)$$

We shall, on this account, say of matrices $\boldsymbol{U}$ with property (70.1) that they are "$\boldsymbol{G}$-unitary." We recover ordinary unitarity by sending $\boldsymbol{G} \to \boldsymbol{I}$. By an easy argument, $\boldsymbol{U} = e^{\boldsymbol{B}}$ will be $\boldsymbol{G}$-unitary if and only if $\boldsymbol{B} \equiv \|B^a{}_b\|$ is

$$\text{"}\boldsymbol{G}\text{-antihermitian"} \quad : \quad \boldsymbol{G}^{-1}\boldsymbol{B}^\dagger\boldsymbol{G} = -\boldsymbol{B}$$

Which (by the hermiticity of $\boldsymbol{G} = \|G_{ab}\|$) entails that $\boldsymbol{G}\boldsymbol{B}$ be antihermitian in the ordinary sense: $(\boldsymbol{G}\boldsymbol{B})^\dagger = -\boldsymbol{G}\boldsymbol{B}$.

A similar (but more familiar) argument establishes that $\boldsymbol{\Lambda} = e^{\boldsymbol{A}}$ will be a Lorentz matrix (i.e., that it will satisfy $\boldsymbol{g} = \boldsymbol{\Lambda}^\intercal\boldsymbol{g}\,\boldsymbol{\Lambda}$) if and only if $\boldsymbol{A} \equiv \|A^\mu{}_\nu\|$ is

$$\text{"}\boldsymbol{g}\text{-antisymmetric"} \quad : \quad \boldsymbol{g}^{-1}\boldsymbol{A}^\intercal\,\boldsymbol{g} = -\boldsymbol{A}$$

Which (by the real symmetry of $\boldsymbol{g} = \|g_{\mu\nu}\|$) entails that $\boldsymbol{g}\boldsymbol{A}$ be antisymmetric in the ordinary sense: $(\boldsymbol{g}\boldsymbol{A})^\intercal = -\boldsymbol{g}\boldsymbol{A}$.

Now write

$$\begin{aligned}\boldsymbol{U} &= \boldsymbol{I} + \boldsymbol{B} + \tfrac{1}{2}\boldsymbol{B}^2 + \cdots\\ \boldsymbol{U}^{-1} &= \boldsymbol{I} - \boldsymbol{B} + \tfrac{1}{2}\boldsymbol{B}^2 - \cdots\\ \boldsymbol{\Lambda} &= \boldsymbol{I} + \boldsymbol{A} + \tfrac{1}{2}\boldsymbol{A}^2 + \cdots\end{aligned}$$

stick those series into (70.2), abandon all but leading order terms, and obtain

$$\boldsymbol{\gamma}^\mu\boldsymbol{B} - \boldsymbol{B}\,\boldsymbol{\gamma}^\mu = A^\mu{}_\nu\boldsymbol{\gamma}^\nu \tag{71}$$

One could, by further cultivation of the algebra, deduce[39]—alternatively: one can, by direct computation, simply confirm—that the solution of (71) can be described

$$\boldsymbol{B} = \tfrac{1}{8}A^{\alpha\beta}(\boldsymbol{\gamma}_\alpha\boldsymbol{\gamma}_\beta - \boldsymbol{\gamma}_\beta\boldsymbol{\gamma}_\alpha) \tag{72}$$

Little of the compact elegance of this result (the neat structure of which might almost have been guessed) survives when it is written out explicitly

$$\boldsymbol{B} = \tfrac{1}{2}\begin{pmatrix} 0 & A_{13} & -A_{03} & -A_{01}\\ -A_{13} & 0 & -A_{01} & A_{03}\\ -A_{03} & -A_{01} & 0 & A_{13}\\ -A_{01} & A_{03} & -A_{13} & 0\end{pmatrix} + i\tfrac{1}{2}\begin{pmatrix} -A_{12} & -A_{23} & 0 & A_{02}\\ -A_{23} & A_{12} & -A_{02} & 0\\ 0 & A_{02} & -A_{12} & -A_{23}\\ -A_{02} & 0 & -A_{23} & A_{12}\end{pmatrix}$$

but the explicit description is not without it uses.[40] For example: Suppose $\boldsymbol{A}$ generates *rotation about the x^3-axis*. We set

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & \vartheta & 0\\ 0 & -\vartheta & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

---

[39] See, for example, Schweber,[7] p. 77.

[40] Note that I have used $A_{\mu\nu} = -A_{\nu\mu}$ to arrange that only the independently specifiable parameters $\{A_{01}, A_{02}, A_{03}, A_{12}, A_{13}, A_{23}\}$ appear in the matrix. It is gratifying to observe that

$$\boldsymbol{G}\boldsymbol{B} = (\text{real antisymmetric}) + i(\text{real symmetric})$$

is in fact antihermitian (as required). Writing $\boldsymbol{B}$ and the explicit descriptions

and (ask *Mathematica* to compute $\sum_0^6 \frac{1}{n!}$`MatrixPower[A,n]//MatrixForm`) are brought to the (anticipated) conclusion that

$$\boldsymbol{\Lambda} = e^{\boldsymbol{A}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\vartheta & \sin\vartheta & 0 \\ 0 & -\sin\vartheta & \cos\vartheta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The generator of the associated spin transformation is

$$\boldsymbol{B} = i\tfrac{1}{2} \begin{pmatrix} -\vartheta & 0 & 0 & 0 \\ 0 & +\vartheta & 0 & 0 \\ 0 & 0 & -\vartheta & 0 \\ 0 & 0 & 0 & +\vartheta \end{pmatrix}$$

which gives

$$\boldsymbol{U} = e^{\boldsymbol{B}} = \begin{pmatrix} e^{-i\frac{1}{2}\vartheta} & 0 & 0 & 0 \\ 0 & e^{+i\frac{1}{2}\vartheta} & 0 & 0 \\ 0 & 0 & e^{-i\frac{1}{2}\vartheta} & 0 \\ 0 & 0 & 0 & e^{+i\frac{1}{2}\vartheta} \end{pmatrix}$$

This result exposes the striking (and deeply consequential) fact that one must rotate 720° in physical space to achieve a 360° rotation in spin space. To describe a *boost in the x^1-direction* we would set

$$\boldsymbol{A} = \begin{pmatrix} 0 & \psi & 0 & 0 \\ \psi & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

and obtain

$$\boldsymbol{\Lambda} = e^{\boldsymbol{A}} = \begin{pmatrix} \cosh\psi & \sinh\psi & 0 & 0 \\ \sinh\psi & \cosh\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

(continued from the preceding page) (54) of $\boldsymbol{\Gamma}^\mu$ into *Mathematica*, it takes only typing time to confirm that

$$\begin{aligned} &\boldsymbol{\Gamma}_0\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_0 - (\ \ A_{01}\boldsymbol{\Gamma}^1 + A_{02}\boldsymbol{\Gamma}^2 + A_{03}\boldsymbol{\Gamma}^3) \\ &= \boldsymbol{\Gamma}_0\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_0 - (-A_{01}\boldsymbol{\Gamma}_1 - A_{02}\boldsymbol{\Gamma}_2 - A_{03}\boldsymbol{\Gamma}_3) = \mathbf{0} \end{aligned}$$

$$\begin{aligned} &\boldsymbol{\Gamma}_1\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_1 - (\ \ A_{10}\boldsymbol{\Gamma}^0 + A_{12}\boldsymbol{\Gamma}^2 + A_{13}\boldsymbol{\Gamma}^3) \\ &= \boldsymbol{\Gamma}_1\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_1 - (-A_{01}\boldsymbol{\Gamma}_0 - A_{12}\boldsymbol{\Gamma}_2 - A_{13}\boldsymbol{\Gamma}_3) = \mathbf{0} \end{aligned}$$

$$\begin{aligned} &\boldsymbol{\Gamma}_2\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_2 - (\ \ A_{20}\boldsymbol{\Gamma}^0 + A_{21}\boldsymbol{\Gamma}^1 + A_{23}\boldsymbol{\Gamma}^3) \\ &= \boldsymbol{\Gamma}_2\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_2 - (-A_{02}\boldsymbol{\Gamma}_0 + A_{12}\boldsymbol{\Gamma}_1 - A_{23}\boldsymbol{\Gamma}_3) = \mathbf{0} \end{aligned}$$

$$\begin{aligned} &\boldsymbol{\Gamma}_3\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_3 - (\ \ A_{30}\boldsymbol{\Gamma}^0 + A_{31}\boldsymbol{\Gamma}^1 + A_{32}\boldsymbol{\Gamma}^2) \\ &= \boldsymbol{\Gamma}_3\boldsymbol{B} - \boldsymbol{B}\boldsymbol{\Gamma}_3 - (-A_{03}\boldsymbol{\Gamma}_0 + A_{13}\boldsymbol{\Gamma}_1 + A_{23}\boldsymbol{\Gamma}_2) = \mathbf{0} \end{aligned}$$

which is the upshot of the assertion that (72) satisfies (71).

The generator of the associated spin transformation is

$$\boldsymbol{B} = \tfrac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -\psi \\ 0 & 0 & -\psi & 0 \\ 0 & -\psi & 0 & 0 \\ -\psi & 0 & 0 & 0 \end{pmatrix}$$

which gives

$$\boldsymbol{U} = e^{\boldsymbol{B}} = \begin{pmatrix} \cosh\tfrac{1}{2}\psi & 0 & 0 & -\sinh\tfrac{1}{2}\psi \\ 0 & \cosh\tfrac{1}{2}\psi & -\sinh\tfrac{1}{2}\psi & 0 \\ 0 & -\sinh\tfrac{1}{2}\psi & \cosh\tfrac{1}{2}\psi & 0 \\ -\sinh\tfrac{1}{2}\psi & 0 & 0 & \cosh\tfrac{1}{2}\psi \end{pmatrix}$$

Contact with the more familiar kinematic parameters β and $\gamma = 1/\sqrt{1-\beta^2}$ is made by means of

$$\begin{aligned} \cosh\psi &= \gamma \\ \sinh\psi &= \gamma\beta \\ \cosh\tfrac{1}{2}\psi &= \sqrt{\tfrac{1}{2}(\gamma+1)} \\ \sinh\tfrac{1}{2}\psi &= \sqrt{\tfrac{1}{2}(\gamma-1)} \end{aligned}$$

Noether meets Einstein. Noether has given us a deep-seated mechanism for associating "current vectors" $J^\mu_r(\varphi, \partial\varphi)$ with parameterized maps (one for each parameter). And for associating conservation laws $\partial_\mu J^\mu_r = 0$ with maps which refer to symmetries of the dynamical action. In specific applications the maps in question are typically contingent—recommended to our attention by features of the dynamical system in hand.

Einstein has directed our attention to a transformation group (the Lorentz group—more generally, the Poincaré group) which the Principle of Relativity asserts to be *universal*, a symmetry shared by *all* admissible dynamical systems, and (in the absence of gravitational effects) explicitly/implicitly present in all well-designed dynamical theories. The stress-energy tensor $S^\mu{}_\nu$ refers to the ν-indexed quartet of conservation laws which arise from the translational part of the Poincaré group. We look now to the design and properties of the sextet of Noetherean currents which arise from the postulated Lorentz covariance of relativistic systems.

Our first assignment is to obtain descriptions of the structure functions characteristic of the homogeneous Lorentz group. For infinitesimal $\boldsymbol{A}$ we have

$$\boldsymbol{\Lambda} = e^{\boldsymbol{A}} = \boldsymbol{I} + \boldsymbol{A} + \cdots$$

$$\boldsymbol{A} = \begin{pmatrix} 0 & \delta\Omega^1 & \delta\Omega^2 & \delta\Omega^3 \\ \delta\Omega^1 & 0 & +\delta\omega^3 & -\delta\omega^2 \\ \delta\Omega^2 & -\delta\omega^3 & 0 & +\delta\omega^1 \\ \delta\Omega^3 & +\delta\omega^2 & -\delta\omega^1 & 0 \end{pmatrix} \equiv \boldsymbol{B}_k\delta\Omega^k + \boldsymbol{R}_k\delta\omega^k$$

where the $\{\boldsymbol{B}_1, \boldsymbol{B}_2, \boldsymbol{B}_3\}$ generate *boosts*, and $\{\boldsymbol{R}_1, \boldsymbol{R}_2, \boldsymbol{R}_3\}$ generate spatial *rotations*. But it is, for the purposes at hand, advantageous to proceed a bit more circumspectly, writing

$$\boldsymbol{A} = \begin{pmatrix} 0 & \delta\Omega_{01} & \delta\Omega_{02} & \delta\Omega_{03} \\ \delta\Omega_{01} & 0 & -\delta\Omega_{12} & +\delta\Omega_{31} \\ \delta\Omega_{02} & +\delta\Omega_{12} & 0 & -\delta\Omega_{23} \\ \delta\Omega_{03} & -\delta\Omega_{31} & +\delta\Omega_{23} & 0 \end{pmatrix}$$

$$\equiv \boldsymbol{N}^{01}\delta\Omega_{01} + \boldsymbol{N}^{02}\delta\Omega_{02} + \boldsymbol{N}^{03}\delta\Omega_{03} + \boldsymbol{N}^{12}\delta\Omega_{12} + \boldsymbol{N}^{23}\delta\Omega_{23} + \boldsymbol{N}^{31}\delta\Omega_{31}$$

$$= \tfrac{1}{2}\boldsymbol{N}^{\alpha\beta}\delta\Omega_{\alpha\beta} \tag{73}$$

subject to the understanding that $\boldsymbol{N}^{\alpha\beta} = -\boldsymbol{N}^{\beta\alpha}$, $\delta\Omega_{\alpha\beta} = -\delta\Omega_{\beta\alpha}$; the generic superscript ^{r} which entered the discussion immediately prior to (1–26) has now become an antisymmetrized *pair* of indices, while the generic $\sum_r$ has become a double sum, managed by the Einstein summation convention. One verifies by inspection that

$$\boldsymbol{N}^{\alpha\beta} = \|N^{\alpha\beta\mu}{}_{\nu}\| = \|(g^{\mu\alpha}\delta^{\beta}{}_{\nu} - g^{\mu\beta}\delta^{\alpha}{}_{\nu})\| \tag{74}$$

We conclude from (74) that the effect, within spacetime, of an infinitesimal Lorentz transformation can be described (see again (1–26.1))

$$\begin{aligned} x^\mu \longrightarrow x^\mu + \delta_\Omega x^\mu & \\ \delta_\Omega x^\mu = \tfrac{1}{2}\mathcal{X}^{\mu\alpha\beta}(x)\delta\Omega_{\alpha\beta} & \\ \mathcal{X}^{\mu\alpha\beta}(x) = (g^{\mu\alpha}\delta^{\beta}{}_{\nu} - g^{\mu\beta}\delta^{\alpha}{}_{\nu})x^\nu & \\ = (g^{\mu\alpha}x^\beta - g^{\mu\beta}x^\alpha) & \end{aligned} \tag{75.1}$$

But at the same time (see again (11–26.2)) the field components $\varphi^a(x)$ fold among themselves in infinitesimal *representation* of the infinitesimal Lorentz transformation; the details are case-specific (depend upon the representation selected by the system in question), but have the generic form

$$\begin{aligned} \varphi^a(x) \longrightarrow \varphi^a(x) + \delta_\Omega\varphi^a(x) & \\ \delta_\Omega\varphi^a(x) = \tfrac{1}{2}\Phi^{a\alpha\beta}(\varphi)\delta\Omega_{\alpha\beta} & \\ \Phi^{a\alpha\beta}(\varphi) = B^a{}_b{}^{\alpha\beta}\varphi^b & \end{aligned} \tag{75.2}$$

If, for example, the physical system involved only a scalar field we would omit the component-identifier (i.e., we would write φ in place of φ^a) and would have $\Phi^{\alpha\beta} = 0$. If the system presented only a single vector field we would in place of φ^a write φ^μ, and would obtain $\Phi^{\mu\alpha\beta}(\varphi) = (g^{\mu\alpha}\varphi^\beta - g^{\mu\beta}\varphi^\alpha)$. We will look later to the relatively more interesting case of the Dirac spinor field, but for the moment proceed generically—armed only with the knowledge that $\Phi^{a\alpha\beta}(\varphi)$ is, like $\mathcal{X}^{\mu\alpha\beta}(x)$, $\alpha\beta$-antisymmetric.

Returning with the structure functions (75) to Noether's equation (1–29), we are led to the $\alpha\beta$-indexed antisymmetric array of currents

$$\begin{aligned} J^{\mu\alpha\beta} &= \frac{\partial \mathcal{L}}{\partial \varphi^a_{,\mu}} \left\{ \Phi^{a\alpha\beta} - \varphi^a_{,\sigma} \mathcal{X}^{\sigma\alpha\beta} \right\} + \mathcal{L} \mathcal{X}^{\mu\alpha\beta} \\ &= - \left\{ \frac{\partial \mathcal{L}}{\partial \varphi^a_{,\mu}} \varphi^a_{,\sigma} - \mathcal{L} \delta^\mu{}_\sigma \right\} \mathcal{X}^{\sigma\alpha\beta} + \frac{\partial \mathcal{L}}{\partial \varphi^a_{,\mu}} \Phi^{a\alpha\beta} \end{aligned}$$

Recalling from (1–34) the construction of the stress-energy tensor,[41] and drawing upon the specific information established at (75.1), we have

$$\begin{aligned} \mathcal{J}^{\mu\alpha\beta} \equiv \tfrac{1}{\mathrm{C}} J^{\mu\alpha\beta} &= -\tfrac{1}{\mathrm{C}} \mathcal{T}^\mu{}_\sigma (g^{\sigma\alpha} x^\beta - g^{\sigma\beta} x^\alpha) + \tfrac{1}{\mathrm{C}} \frac{\partial \mathcal{L}}{\partial \varphi^a_{,\mu}} \Phi^{a\alpha\beta} \\ &= \tfrac{1}{\mathrm{C}} (x^\alpha \mathcal{T}^{\mu\beta} - x^\beta \mathcal{T}^{\mu\alpha}) + \tfrac{1}{\mathrm{C}} \frac{\partial \mathcal{L}}{\partial \varphi^a_{,\mu}} B^a{}_b{}^{\alpha\beta} \varphi^b \\ &= \mathcal{M}^{\mu\alpha\beta} + S^{\mu\alpha\beta} \end{aligned} \tag{76}$$

where, in terminology which it will be incumbent upon us to justify,

$$\mathcal{M}^{\mu\alpha\beta} \equiv \tfrac{1}{\mathrm{C}} (x^\alpha \mathcal{T}^{\mu\beta} - x^\beta \mathcal{T}^{\mu\alpha}) \;:\; \begin{cases} \text{defines the "orbital angular momentum"} \\ \text{of the relativistic field system} \end{cases}$$

$$S^{\mu\alpha\beta} \equiv \tfrac{1}{\mathrm{C}} \frac{\partial \mathcal{L}}{\partial \varphi^a_{,\mu}} B^a{}_b{}^{\alpha\beta} \varphi^b \;:\; \begin{cases} \text{defines the "intrinsic angular momentum"} \\ \text{or "spin" of the system} \end{cases}$$

The $\frac{1}{\mathrm{C}}$-factors have been introduced to render such terminology *dimensionally* tenable:

$$[\,\mathcal{J}^{\mu\alpha\beta}\,] = [\,\mathcal{M}^{\mu\alpha\beta}\,] = [\,S^{\mu\alpha\beta}\,] = \frac{(\text{length})(\text{energy density})}{\text{velocity}}$$
$$= \text{angular momentum density}$$

The expressions $\mathcal{J}^{\mu\alpha\beta}(\varphi, \partial\varphi, x)$ (ditto $\mathcal{M}^{\mu\alpha\beta}(\varphi, \partial\varphi, x)$, $S^{\mu\alpha\beta}(\varphi)$) transform as components of an $\alpha\beta$-antisymmetric third-rank tensor. The components of $\mathcal{J}^{\mu\alpha\beta}$ (ditto $\mathcal{M}^{\mu\alpha\beta}$, $S^{\mu\alpha\beta}$) are intermixed by Lorentz transformation, but the resolution of $\mathcal{J}^{\mu\alpha\beta}(\varphi, \partial\varphi, x)$ into an "orbital" part and an "intrinsic" part ("spin") maintains its integrity (i.e., is something all inertial observers agree upon).

From the Lorentz invariance which was built into the Lagrangian (whence into the action functional) we know it to be an implication of the (Lorentz covariant) field equations that

$$\partial_\mu \mathcal{J}^{\mu\alpha\beta} = 0 \tag{77}$$

[41] For the purpose of this discussion (in order to release the letter S to *spin*-like assignments) we use $\mathcal{T}^{\mu\nu}$ to denote the stress-energy tensor (which many authors call the "energy-momentum tensor.")

This is a *sextet of conservation laws, which pertain with certainty to every relativistic field theory*. Drawing motivation from this result and from the additive structure of (76), we observe that

$$\begin{aligned}\partial_\mu \mathcal{M}^{\mu\alpha\beta} &= \tfrac{1}{\mathrm{c}}\Big\{\delta^\alpha{}_\mu \mathcal{T}^{\mu\beta} + x^\alpha \partial_\mu \mathcal{T}^{\mu\beta} - \delta^\beta{}_\mu \mathcal{T}^{\mu\alpha} + x^\beta \partial_\mu \mathcal{T}^{\mu\alpha}\Big\} \\ &= \tfrac{1}{\mathrm{c}}\Big\{\mathcal{T}^{\alpha\beta} - \mathcal{T}^{\beta\alpha}\Big\} \quad \text{by } \partial_\mu \mathcal{T}^{\mu\nu} = 0 \\ &= 0 \quad \text{if and only if the stress-energy tensor } \mathcal{T}^{\mu\nu} \text{ is } \textit{symmetric}\end{aligned}$$

Evidently (77) is, in non-symmetric cases, achieved by a kind of "trade-off" between the orbital and intrinsic parts of $\mathcal{J}^{\mu\alpha\beta}$. Such "spin-orbit coupling" is, of course, commonplace in the classical mechanics of many-body systems (tops, planetary systems), as it is also in atomic physics.

Look now to the expressions

$$\begin{aligned}\mathcal{M}^{\alpha\beta} &\equiv \iiint \mathcal{M}^{0\alpha\beta}\, dx^1 dx^2 dx^3 \\ &= \tfrac{1}{\mathrm{c}} \iiint \left(x^\alpha \mathcal{T}^{0\beta} - x^\beta \mathcal{T}^{0\alpha}\right) dx^1 dx^2 dx^3\end{aligned}$$

which evidently refer to *first moment properties* of the expressions

$$\mathcal{T}^{0\nu} \equiv c\mathcal{P}^\nu \;:\; \begin{cases} c\mathcal{P}^0 = \mathcal{E} & : \quad \text{energy density} \\ \mathcal{P}^1, \mathcal{P}^2, \mathcal{P}^3 & : \quad \text{components of momentum density} \end{cases}$$

and which in symmetric cases $\mathcal{T}^{\mu\nu} = \mathcal{T}^{\nu\mu}$ will describe global *constants of the field motion* (they will in other cases remain "interesting," even though not conserved). It is an option available to each individual inertial observer to resolve those expressions into two classes:

$$\begin{aligned} K_1 &\equiv \mathcal{M}^{01} & \qquad L_1 &\equiv \mathcal{M}^{23} \\ K_2 &\equiv \mathcal{M}^{02} & \qquad L_2 &\equiv \mathcal{M}^{31} \\ K_3 &\equiv \mathcal{M}^{03} & \qquad L_3 &\equiv \mathcal{M}^{12} \end{aligned}$$

Looking first to the latter, we have

$$\begin{aligned} L_1 &= \iiint \left(x^2 \mathcal{P}^3 - x^3 \mathcal{P}^2\right) dx^1 dx^2 dx^3 \\ L_2 &= \iiint \left(x^3 \mathcal{P}^1 - x^1 \mathcal{P}^3\right) dx^1 dx^2 dx^3 \\ L_3 &= \iiint \left(x^1 \mathcal{P}^2 - x^2 \mathcal{P}^1\right) dx^1 dx^2 dx^3 \end{aligned}$$

which lend a clear and natural meaning to the *orbital angular momentum density* $\boldsymbol{\mathcal{L}} \equiv \boldsymbol{x} \times \boldsymbol{\mathcal{P}}$ of a relativistic field system. An identical train of thought

leads one to write

$$K_1 = c\iiint \left(t\mathcal{P}^1 - x^1\mathcal{M}\right) dx^1dx^2dx^3$$
$$K_2 = c\iiint \left(t\mathcal{P}^2 - x^2\mathcal{M}\right) dx^1dx^2dx^3$$
$$K_3 = c\iiint \left(t\mathcal{P}^3 - x^3\mathcal{M}\right) dx^1dx^2dx^3$$

where we have used $x^0 = ct$ and introduced

$$\mathcal{M} \equiv \mathcal{E}/c^2 \ : \ \text{equivalent "mass density"}$$

Write

$$M \equiv \iiint \mathcal{M}\, dx^1dx^2dx^3 = \text{conserved "total mass" of the field system}$$
$$= (\text{conserved total energy})/c^2$$

$$\boldsymbol{X} \equiv \tfrac{1}{M}\iiint \boldsymbol{x}\,\mathcal{M}\, dx^1dx^2dx^3 = \text{"center of mass" of the field system}$$

$$\boldsymbol{P} \equiv \iiint \boldsymbol{\mathcal{P}}\, dx^1dx^2dx^3 = \text{conserved total momentum}$$

and obtain the conservation of $\boldsymbol{K} = t\boldsymbol{P} - M\boldsymbol{X}$, which can be expressed

$$\boldsymbol{X}(t) = \tfrac{1}{M}\boldsymbol{P}\,t - \boldsymbol{K}$$

We are brought thus to the conclusion that *if the stress-energy tensor* $\mathcal{T}^{\mu\nu}$ *is symmetric* (!) then

- orbital angular momentum $\boldsymbol{L} \equiv \iiint \boldsymbol{\mathcal{L}}\, dx^1dx^2dx^3$ is conserved;
- the center of mass $\boldsymbol{X}(t)$ of the field system moves uniformly/rectilinearly;
- six conservation laws $\partial_\mu S^{\mu\alpha\beta} = 0$ refer to the "conservation of spin."

Belinfante's fandango. The pretty results obtained in discussion subsequent to (77) are, as I have several times stressed, contingent upon the *symmetry of the stress-energy tensor*. But introduction of (76) into (77) gives

$$\mathcal{T}^{\alpha\beta} - \mathcal{T}^{\beta\alpha} + c\,\partial_\mu S^{\mu\alpha\beta} = 0 \tag{78}$$

so in point of fact we have $\mathcal{T}^{\alpha\beta} = \mathcal{T}^{\beta\alpha}$ only with respect to field components φ^a and φ^b which transform as *scalars*.[42] However...

Noetherian currents generally (and the stress-energy tensor in particular) respond non-trivially to unphysical gauge transformations:[43]

$$\mathcal{L} \longrightarrow L = \mathcal{L} + \partial_\sigma \mathcal{G}^\sigma(\varphi) \tag{79.1}$$

[42] Such fields are too simple to respond to Lorentz transformation by "folding among themselves," and $B^a{}_b{}^{\alpha\beta} = 0$ forces $\mathcal{S}^{\mu\alpha\beta} = 0$.

[43] Here I simply adapt to the language of *relativistic* field theory (and extend) the remarks which accompanied (1–36).

induces

$$\mathcal{J}_r^\mu \longrightarrow J_r^\mu = \mathcal{J}_r^\mu + G_r^\mu(\varphi, \partial\varphi) \tag{79.2}$$
$$G_r^\mu(\varphi, \partial\varphi) = \left\{ \left[\Phi_{ar} - \varphi_{a,\rho}\mathcal{X}_r^\rho\right] \frac{\partial}{\partial \varphi_{a,\mu}} + \mathcal{X}_r^\mu \right\} \left(\partial_\sigma \mathcal{G}^\sigma\right)$$

so one must be circumspect when ascribing "direct physical significance" to properties of such currents. In relativistic field theory it is natural to require of $\mathfrak{L} \longrightarrow L$ that it respect the "principle of Lagrangian Lorentz-invariance," which is automatic if one admits only expressions $\mathcal{G}^\sigma(\varphi)$ which transform as 4-vectors. The gauge transformation will "preserve conservation laws"

$$\partial_\mu \mathcal{J}_r^\mu = 0 \quad \Longrightarrow \quad \partial_\mu J_r^\mu = 0 \quad \text{if and only if } \partial_\mu G_r^\mu = 0 \tag{80.1}$$

and the latter condition becomes automatic in cases where G_r^ν can be described

$$G_r^\nu(\varphi, \partial\varphi) = \partial_\mu H_r^{\mu\nu}(\varphi) \quad \text{with} \quad H_r^{\mu\nu} = -H_r^{\nu\mu} \tag{80.2}$$

From

$$\begin{aligned}
\mathcal{J}_r &= \iiint_{\mathcal{R}} \mathcal{J}_r^0 \, dx^1 dx^2 dx^3 \\
&= \iiint_{\mathcal{R}} J_r^0 \, dx^1 dx^2 dx^3 + \iiint_{\mathcal{R}} \boldsymbol{\nabla}\cdot\boldsymbol{H}_r \, dx^1 dx^2 dx^3 \quad : \quad \text{here } \boldsymbol{H}_r \equiv \begin{pmatrix} H_r^{10} \\ H_r^{20} \\ H_r^{30} \end{pmatrix} \\
&= \iiint_{\mathcal{R}} J_r^0 \, dx^1 dx^2 dx^3 + \iint_{\partial\mathcal{R}} \boldsymbol{H}_r \cdot d\boldsymbol{\sigma} \\
&= J_r \quad \text{on the presumption that the surface integral vanishes as } \partial\mathcal{R} \to \infty
\end{aligned}$$

we see that $\partial_\mu \mathcal{J}_r^\mu = 0$ and $\partial_\mu J_r^\mu = 0$ will then refer in distinct local dialects to the global conservation of the same net things…to which, however, they ascribe distinct densities $\mathcal{J}_r^0 \neq J_r^0$ and correspondingly distinct fluxes $\boldsymbol{\mathcal{J}}_r \neq \boldsymbol{J}_r$.

Look now, within the framework of those remarks, to the stress-energy (which arises from the translational part of the Poincaré group; the generic $_r$ becomes now the Greek index $_\nu$). We have

$$\mathcal{T}^\mu{}_\nu \longrightarrow T^\mu{}_\nu = \mathcal{T}^\mu{}_\nu + \partial_\sigma H^{\sigma\mu}{}_\nu \tag{81}$$

If $\mathfrak{L} \longrightarrow L$ preserves translational invariance then it will assuredly be the case that both $\partial_\mu \mathcal{T}^\mu{}_\nu = 0$ and $\partial_\mu T^\mu{}_\nu = 0$; in the preceding equation the latter statement is presented as a corollary of the former, attributable to the $\sigma\mu$-antisymmetry of $H^{\sigma\mu}{}_\nu$.

The idea now—original to F. J. Belinfante[44]—is to use gauge freedom to achieve the symmetry of $T^{\mu\nu}$. We proceed from the observation that

$$T^{\mu\nu} - T^{\nu\mu} = 0 \quad \text{entails} \quad \mathcal{T}^{\mu\nu} - \mathcal{T}^{\nu\mu} + \partial_\sigma(H^{\sigma\mu\nu} - H^{\sigma\nu\mu}) = 0$$

[44] "On the spin angular momentum of mesons," Physica **7**, 882 (1939).

which by (78) becomes

$$\partial_\sigma(H^{\sigma\mu\nu} - H^{\sigma\nu\mu}) = c\,\partial_\sigma S^{\sigma\mu\nu}$$

These derivative conditions would certainly be satisfied if $H^{\sigma\mu\nu}$ satisfied the algebraic conditions

$$H^{\sigma\mu\nu} - H^{\sigma\nu\mu} = c\,S^{\sigma\mu\nu} \tag{82}$$

But $H^{\sigma\nu\mu} = -H^{\nu\sigma\mu}$ so we have the first of the following equations (the other two equations—redundant with the first—are got by cyclic permutation of indices):

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} H^{\mu\nu\sigma} \\ H^{\sigma\mu\nu} \\ H^{\nu\sigma\mu} \end{pmatrix} = c \begin{pmatrix} S^{\sigma\mu\nu} \\ S^{\nu\sigma\mu} \\ S^{\mu\nu\sigma} \end{pmatrix}$$

By matrix inversion we obtain

$$\begin{pmatrix} H^{\mu\nu\sigma} \\ H^{\sigma\mu\nu} \\ H^{\nu\sigma\mu} \end{pmatrix} = \tfrac{1}{2}c \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} S^{\sigma\mu\nu} \\ S^{\nu\sigma\mu} \\ S^{\mu\nu\sigma} \end{pmatrix}$$

Thus by a little trickery (the "permutation trick," which we will have occasion to use again in quite another connection) do we obtain

$$H^{\sigma\mu\nu} = \tfrac{1}{2}c\big\{S^{\sigma\mu\nu} + S^{\mu\nu\sigma} - S^{\nu\sigma\mu}\big\} \tag{83}$$

Since $S^{\mu\alpha\beta} \equiv \frac{1}{\mathfrak{C}}\pi^\mu_a B^a{}_b{}^{\alpha\beta}\varphi^b$ (here $\pi^\mu_a \equiv \partial\mathfrak{L}/\partial\varphi^a_{,\mu}$) is antisymmetric in the last pair of superscripts, the expression on the right can be written (and in the literature is written) in a variety of alternative ways; from $S^{\sigma\mu\nu} = -S^{\sigma\nu\mu}$ it follows in particular that $H^{\sigma\mu\nu} = -H^{\mu\sigma\nu}$, as was stipulated at (80.2).

Returning with (83) to (81) we have Belinfante's

$$\begin{aligned} \mathcal{T}^{\mu\nu} \longrightarrow T^{\mu\nu} &= \mathcal{T}^{\mu\nu} + \vartheta^{\mu\nu} \\ \vartheta^{\mu\nu} &\equiv \partial_\sigma H^{\sigma\mu\nu} \end{aligned} \tag{84}$$

and look now to what that adjustment does to the angular momentum tensor: we have

$$\begin{aligned} \mathcal{J}^{\mu\alpha\beta} &= \tfrac{1}{\mathfrak{C}}(x^\alpha \mathcal{T}^{\mu\beta} - x^\beta \mathcal{T}^{\mu\alpha}) + S^{\mu\alpha\beta} \\ &= \tfrac{1}{\mathfrak{C}}(x^\alpha T^{\mu\beta} - x^\beta T^{\mu\alpha}) - \tfrac{1}{\mathfrak{C}}\underbrace{(x^\alpha\,\vartheta^{\mu\beta} - x^\beta\,\vartheta^{\mu\alpha})} + S^{\mu\alpha\beta} \\ &\qquad\qquad \Big\lceil \\ &\qquad\qquad = x^\alpha \partial_\sigma H^{\sigma\mu\beta} - x^\beta \partial_\sigma H^{\sigma\mu\alpha} \\ &\qquad\qquad = \partial_\sigma(x^\alpha H^{\sigma\mu\beta} - x^\beta H^{\sigma\mu\alpha}) - \underbrace{(H^{\alpha\mu\beta} - H^{\beta\mu\alpha})} \\ &\qquad\qquad\qquad\qquad = -(H^{\mu\alpha\beta} - H^{\mu\beta\alpha}) \\ &\qquad\qquad\qquad\qquad = -c\,S^{\mu\alpha\beta} \quad \text{by (82)} \end{aligned}$$

which (notice that the $S^{\mu\alpha\beta}$-term drops away) can be expressed

$$\mathcal{J}^{\mu\alpha\beta} \longrightarrow J^{\mu\alpha\beta} = \mathcal{J}^{\mu\alpha\beta} + \partial_\sigma(x^\alpha H^{\sigma\mu\beta} - x^\beta H^{\sigma\mu\alpha}) \tag{85}$$

The statements $\partial_\mu \mathcal{J}^{\mu\alpha\beta} = 0$ and $\partial_\mu J^{\mu\alpha\beta} = 0$ are now clearly equivalent, and the latter can be looked upon as a corollary (by $T^{\mu\nu} = T^{\nu\mu}$) of

$$J^{\mu\alpha\beta} \equiv \tfrac{1}{\mathfrak{C}}(x^\alpha\, T^{\mu\beta} - x^\beta\, T^{\mu\alpha}) \tag{86}$$

Evidently $\partial_\mu \mathcal{J}^{\mu\alpha\beta} = 0$ and $\partial_\mu J^{\mu\alpha\beta} = 0$ speak in distinct local dialects about the global conservation of the same six things

$$\iiint \mathcal{J}^{0\alpha\beta}\, dx^1 dx^2 dx^3 = \iiint J^{0\alpha\beta}\, dx^1 dx^2 dx^3 \tag{87.1}$$

as do $\partial_\mu \mathcal{T}^{\mu\alpha} = 0$ and $\partial_\mu T^{\mu\alpha\beta} = 0$ speak in distinct local dialects about the global conservation of the same four things

$$\iiint \mathcal{T}^{0\alpha}\, dx^1 dx^2 dx^3 = \iiint T^{0\alpha}\, dx^1 dx^2 dx^3 \tag{87.2}$$

Notice that the previous definition $\mathcal{M}^{\mu\alpha\beta} \equiv \frac{1}{\mathfrak{C}}(x^\alpha \mathcal{T}^{\mu\beta} - x^\beta \mathcal{T}^{\mu\alpha})$ of "orbital angular momentum" displayed momental structure, but contained no reference to the internal transformation (or spin) properties of the field system. Spin structure (see again (84)) was, however, folded into the design of $T^{\mu\nu}$, and at (86) was incorporated also into the momental design of $J^{\mu\alpha\beta}$. Introduction of (84) into (86) gives

$$\begin{aligned} J^{\mu\alpha\beta} =& M^{\mu\alpha\beta} + S^{\mu\alpha\beta} \\ M^{\mu\alpha\beta} \equiv& \mathcal{M}^{\mu\alpha\beta} + \tfrac{1}{\mathfrak{C}}\partial_\sigma(x^\alpha H^{\sigma\mu\beta} - x^\beta H^{\sigma\mu\alpha}) \end{aligned} \tag{88}$$

Belinfante's procedure[45] leaves the spin tensor $S^{\mu\alpha\beta}$ unchanged, but at the cost

$$\mathcal{M}^{\mu\alpha\beta} \longrightarrow M^{\mu\alpha\beta} = \mathcal{M}^{\mu\alpha\beta} + \tfrac{1}{\mathfrak{C}}\partial_\sigma(x^\alpha H^{\sigma\mu\beta} - x^\beta H^{\sigma\mu\alpha}) \tag{89}$$

of mixing some spin structure into the orbital angular momentum tensor; this, however, has been accomplished in such a way as to leave the (generally not conserved) global expressions

$$\iiint \mathcal{M}^{0\alpha\beta}\, dx^1 dx^2 dx^3$$

[45] The procedure is known in some circles as "symmetrization," in others as "regularization," and is frequently employed without attribution to Belinfante. The symmetrized (or regularized) stress energy tensor $T^{\mu\nu}$ is, by some fastidious authors who wish to underscore the fact that it incorporates spin structure, occasionally called the "spin-stress-energy tensor."

unaltered.

It might seem fair to ask—though the question remains unasked in the literature known to me[46]—for the $\mathcal{G}^\sigma(\varphi)$ which when introduced into (79.1) would set Belinfante's train of argument into prefigured motion. I do not know the answer, or even whether such a $\mathcal{G}^\sigma(\varphi)$ exists, but do not pursue the matter because I think deeper insight into the essence of Belinfante's procedure is, in fact, to be found elsewhere. Heretofore we have presumed that

$$\begin{aligned} x^\mu &\longrightarrow x^\mu + \delta\omega^\mu \\ \varphi_a &\longrightarrow \varphi_a \end{aligned}$$

serves to describe the translation map in spacetime, and have by Noether's theorem been led from this presumption to the familiar stress-energy tensor[47]

$$\mathcal{T}^\mu{}_\nu = \frac{\partial \mathcal{L}}{\partial \varphi_{a,\mu}} \varphi_{a,\nu} - \mathcal{L}\delta^\mu{}_\nu$$

Suppose, however, we were—in the spirit of (1–37)—to adopt this expanded conception of "translation:"

$$\left.\begin{aligned} x^\mu &\longrightarrow x^\mu + g^{\mu\rho}\cdot\delta\omega_\rho \\ \varphi_a &\longrightarrow \varphi_a \\ \mathcal{L} &\longrightarrow \mathcal{L} - \partial_\sigma \mathcal{G}^{\sigma\rho}(\varphi)\cdot\delta\omega_\rho \end{aligned}\right\} \tag{90}$$

Noether's theorem, correspondingly expanded, would then according to (1–38) supply

$$T^{\mu\nu} = \mathcal{T}^{\mu\nu} + \mathcal{G}^{\mu\nu}$$

which would reproduce (84) if we were to set

$$\mathcal{G}^{\mu\nu} = \vartheta^{\mu\nu} \equiv \partial_\tau H^{\tau\mu\nu}$$

We are led thus from the generic (90) to what might be called "Belinfante's translation map"

$$\left.\begin{aligned} x^\mu &\longrightarrow x^\mu + g^{\mu\rho}\cdot\delta\omega_\rho \\ \varphi_a &\longrightarrow \varphi_a \\ \mathcal{L} &\longrightarrow \mathcal{L} + \partial_\sigma\partial_\tau H^{\sigma\tau\rho}\cdot\delta\omega_\rho \end{aligned}\right\} \tag{91}$$

[46] My sources have been W. Pauli, "Relativistic field theories of elementary particles," Rev. Mod. Phys. **13**, 203 (1941), §2; E. M. Corson, *Introduction to Tensors, Spinors, and Relativistic Wave Equations* (1953), §19—Corson, by the way, combined careers in theoretical/mathematical physics and psychiatry! —and J. M. Jauch's appendix to G. Wentzel, *Quantum Theory of Fields* (1949). More recently the subject has been treated with expository freshness by D. E. Soper, *Classical Field Theory* (1976), pp. 116–123.

[47] See again (1–34).

which can be looked upon as the seed from which Belinfante's symmetrization procedure—in its entirety—spontaneously springs. Here $H^{\sigma\tau\rho}$ is understood to have the specific meaning indicated at (83), which can be written

$$H^{\sigma\tau\rho} = \tfrac{1}{2}\big\{\pi_a^\sigma B^a{}_b{}^{\tau\rho} + \pi_a^\tau B^a{}_b{}^{\rho\sigma} - \pi_a^\rho B^a{}_b{}^{\sigma\tau}\big\}\varphi^b \tag{92}$$

We note in passing that the gauge terms

$$\partial_\sigma\partial_\tau H^{\sigma\tau\rho} = 0 \quad \text{numerically}$$

and that (92) is in one (slight?) respect problematic.[48]

The intrusion of a gauge component into the definition of a map is certainly not unprecedented,[49] but the occurance of an allusion to *spin structure* in the definition of an infinitesimal *translation* is a bit of a surprise; the translational and Lorentzian parts of the Poincaré group have in (91) become fused.

Belinfante's procedure works for *any* relativistic Lagrangian field theory. It owes its success, however, not so much to relativity *per se* as to the spin structure which relativity brings to the physics. Davison Soper[46] has described a non-relativistic analog of Belinfante's procedure which again draws upon spin structure, but presumes only *rotational* invariance of the Lagrangian.

What—if any—*physical* significance are we to associate with the distinction between

$$\mathcal{T}^{\mu\alpha} \quad \text{and} \quad T^{\mu\alpha} = \mathcal{T}^{\mu\alpha} + (\text{divergenceless gauge term})^{\mu\alpha} \tag{93}$$

It is sometimes argued in electrodynamics that the field $F^{\mu\nu}$ announces itself never directly, but always indirectly—through the associated stress-energy tensor, which mediates the interaction of field with charged matter. But when one looks to the equation which describes that interaction[50] one observes that it is of a form

$$\partial_\mu(\text{stress-energy of field})^{\mu\nu} + \partial_\mu(\text{stress-energy of matter})^{\mu\nu} = 0$$

---

[48] Looking into the abbreviation $\pi_a^\mu \equiv \partial\mathcal{L}/\partial\varphi^a_{,\mu} = \pi_a^\mu(\varphi, \partial\varphi)$, we note that $\partial\varphi$-dependence is unwelcome in a gauge term. The problem does not arise, however, in *canonically* formulated theories (for the same reason that they do not admit of Hamiltonian formulation), and I must on this occasion be content to let it drop. But look to the footnote on p. 121 of Soper.[46]

[49] Construction of a Lagrangian account of the Galilean-covariant systems of particles is found to entail incorporation of a gauge transformation into the definition of a Galilean transformation, and in the case $L = \frac{1}{2}m\dot{x}^2$ leads to a conservation law formally identical to the previously encountered conservation of $\boldsymbol{K} = t\boldsymbol{P} - M\boldsymbol{X}$; see CLASSICAL MECHANICS (1983), p. 169. A similar detail intrudes into the quantum theory of such systems.

[50] See CLASSICAL ELECTRODYNAMICS (1980), p. 312.

into which the stress-energy tensor enters not nakedly, but differentiated. In such contexts the distinction (93) is without consequence. We conclude that transformations of the form

$$\mathcal{J}^\mu_r \longrightarrow J^\mu_r = \mathcal{J}^\mu_r + (\text{divergenceless gauge term})^\mu_r$$

can be considered "physical"—and thus to present a selection problem—only in contexts (should any present themselves) into which the Noetherean current in question enters "nakedly." General relativity presents a context into which the stress-energy tensor does just that, for it is *via* its stress-energy tensor—the repository of information concerning the *distribution* of field-energy/mass—that a field system generates/feels gravitational fields. In that context "the dialect makes a difference." Several lines of argument indicate that it is the symmetrized stress-energy tensor $T^{\mu\nu}$—not $\mathcal{T}^{\mu\nu}$—which speaks the preferred dialect when stripped naked by gravity.[51]

Belinfante's brilliant little mathematical dance is, in my opinion, classic; it uses the simplest and most general of means to achieve a far-reaching—and conceptually challenging—result of great practical importance. Because he had a local history, I allow myself this sentimental digression: Frederic Jozef Belinfante was born in 1913 in The Hague. He took his doctorate in 1939 from the University of Leiden, where his research was directed by H. A. Kramers, and it is (I infer) in his dissertation that the symmetrization procedure—the work of a 26-year-old—was first described; the thesis itself presented "the germ of the connection between the 'spin-statistics theorem' and the yet undiscovered TCP theorem,"[52] and is for that reason still highly regarded. Belinfante was professionally inactive during the war years (which he spent in Holland), but in 1946 accepted a faculty appointment to the University of British Columbia, and in 1948 went to Purdue, from which he retired in 1979. Belinfante had wide-ranging interests outside of physics; he was, like many of his countrymen, a linguist (wrote scientific papers in Esperanto, and introduced—in addition to a great many failed neologisms—the word "nucleon" into the vocabulary of physics), and found endless fascination in maps, railway schedules, meteorological charts. It was study of the latter which led him in his retirement to Portland, which by his calculation presented the best American approximation to the climate of Holland. When in the early 1980's Jean Delord brought to my notice a letter to the editors of *Physics Today* signed by one "F. J. Belinfante, Gresham, Oregon" I immediately looked the man up in the phone book, and gave him a call. Soon I was able to arrange for him to be named Adjunct Professor (Reed's first) and provided with an office, a computer (his

[51] See Léon Rosenfeld, "Sur le tenseur d'impulse-energie," Memoires de l'Academie Roy. Belgique **6**, 30 (1940). Also C. W. Misner, K. S. Thorne & J. A. Wheeler, *Gravitation* (1973), §21.3.

[52] I quote from I. Duck & E. C. G. Sudarshan, *Pauli and the Spin-Statistics Theorem* (1997), p. 301. Chapter 13 of that monograph treats "Belinfante's Proof of the Spin-Statistics Theorem." Belinfante's work led first to a dispute—but later to a collaborative paper (1940)—with Pauli.

first, though when he arrived in Portland he had wasted no time in joining the local HP–45 Club), and people to talk to. Throughout the 1980's he was seen—a slight, grandfatherly man in a large coat, trailing his briefcase on a little cart (a tradition now carried forward by me)—as he made his way from the East Parking Lot to the Physics Building. He gave us some seminars, consumed many of David Griffiths' afternoons discussing computational problems at the blackboard (he loved to talk, had lots of complicated stuff on his mind, and was oblivious of the possibility that David might have other things to do), and busied himself with the simultaneous preparation of two projected books in the tradition of his *A Survey of Hidden-Variable Theories* (1973); one of those was to treat quantum field theory, the other general relativity, and both remained unfinished at the time of his death, in the early summer of 1991. He left his personal library to Reed College. We talked of many things on many occasions. I regret that I never talked with him about the circumstances which led to the invention of his symmetrization procedure.[53]

Conservation laws for some illustrative field systems. My objective here will be to show how the general results developed in the preceding two sections look when brought to ground in particular cases. The conservation laws we will be looking at all derive from the postulated *Poincaré covariance* of the field systems in question.

REAL SCALAR FIELD

Here some of the main results lie already at hand;[54] I write them out again to place them in their larger context, to establish some notation, and to set the pattern to which I will adhere when discussing more complicated systems. The Lagrangian reads

$$\mathfrak{L} = \tfrac{1}{2}\big\{g^{\rho\sigma}\varphi_{,\rho}\varphi_{,\sigma} - \varkappa^2\varphi^2\big\} \qquad (11 \equiv 94.0)$$

The stress-energy tensor derives from the translational component of the Poincaré group; infinitesimally

$$\left.\begin{array}{lll} x^\mu \longrightarrow x^\mu + \mathcal{X}^{\mu\alpha}\cdot\delta\omega_\alpha & \text{with} & \mathcal{X}^{\mu\alpha} = g^{\mu\alpha} \\ \varphi \longrightarrow \varphi + \Phi^\alpha\cdot\delta\omega_\alpha & \text{with} & \Phi^\alpha = 0 \end{array}\right\} \qquad (94.1)$$

which when fed into the generic

$$\begin{aligned} J^{\mu\alpha} =& \pi^\mu\big\{\Phi^\alpha - \varphi_{,\sigma}\mathcal{X}^{\sigma\alpha}\big\} + \mathfrak{L}\mathcal{X}^{\mu\alpha} \qquad (94.2)\\ & \pi^\mu \equiv \partial\mathfrak{L}/\partial\varphi_{,\mu} \end{aligned}$$

[53] For further information see the obituary printed on p. 82 of the July 1992 issue of *Physics Today*. Related material can also be found in M. Dresden, *H. A. Kramers: Between Tradition and Revolution* (1987).

[54] See again the material subsequent to (11). Also (1–41). The funny equation numbers are synchronized to those encountered in subsequent examples.

gives (after the conventional sign reversal)

$$\begin{aligned}\mathcal{T}^{\mu\alpha} &= -\Big[g^{\mu\rho}\varphi_{,\rho}\big\{0-\varphi_{,\sigma}g^{\sigma\alpha}\big\}+\mathcal{L}g^{\mu\alpha}\Big]\\ &= \pi^{\mu}\pi^{\alpha}-\tfrac{1}{2}\big(g_{\rho\sigma}\pi^{\rho}\pi^{\sigma}-\varkappa^{2}\varphi^{2}\big)g^{\mu\alpha}\end{aligned}\tag{94.10}$$

where $\pi^{\mu}=\partial^{\mu}\varphi$ in the case at hand. The stress-energy tensor is already symmetric

$$\mathcal{T}^{\mu\alpha}=\mathcal{T}^{\alpha\mu}\tag{94.11}$$

and the energy density is given by

$$\mathcal{E}=\mathcal{T}^{00}=\tfrac{1}{2}\big\{\pi^{0}\pi^{0}+\pi^{1}\pi^{1}+\pi^{2}\pi^{2}+\pi^{3}\pi^{3}+\varkappa^{2}\varphi^{2}\big\}\geqslant 0\tag{94.12}$$

The (proper) Lorentz component of the Poincaré group gives[55]

$$\left.\begin{aligned}x^{\mu}&\longrightarrow x^{\mu}+\tfrac{1}{2}\mathcal{X}^{\mu\alpha\beta}\cdot\delta\Omega_{\alpha\beta}\quad\text{with}\quad\mathcal{X}^{\mu\alpha\beta}=(g^{\mu\alpha}x^{\beta}-g^{\mu\beta}x^{\alpha})\\ \varphi&\longrightarrow\varphi\ +\tfrac{1}{2}\ \Phi^{\alpha\beta}\cdot\delta\Omega_{\alpha\beta}\quad\text{with}\quad\Phi^{\alpha\beta}=0\end{aligned}\right\}\tag{94.4}$$

which when fed into the generic

$$J^{\mu\alpha\beta}=\pi^{\mu}\big\{\Phi^{\alpha\beta}-\varphi_{,\sigma}\mathcal{X}^{\sigma\alpha\beta}\big\}+\mathcal{L}\mathcal{X}^{\mu\alpha\beta}\tag{94.5}$$

give (compare (76))

$$\begin{aligned}\mathcal{J}^{\mu\alpha\beta}&=\tfrac{1}{\mathrm{c}}\Big[\big(-\pi^{\mu}\pi^{\alpha}+\mathcal{L}g^{\mu\alpha}\big)x^{\beta}-\big(-\pi^{\mu}\pi^{\beta}+\mathcal{L}g^{\mu\beta}\big)x^{\alpha}\Big]\\ &=\tfrac{1}{\mathrm{c}}\big(x^{\alpha}\mathcal{T}^{\mu\beta}-x^{\beta}\mathcal{T}^{\mu\alpha}\big)+\text{no spin component}\end{aligned}\tag{94.16}$$

In the absence of a spin component Belinfante's procedure has nothing to work with, but on the other hand it has nothing to do: stress-energy symmetry ensures

$$\partial_{\mu}\mathcal{J}^{\mu\alpha\beta}=0$$

Similar results are obtain in the case

$$\mathcal{L}=\tfrac{1}{2}\big\{g^{\rho\sigma}\psi^{*}_{,\rho}\psi_{,\sigma}-\varkappa^{2}\psi^{*}\psi\big\}\tag{19}$$

of a complex scalar field.

PROCA THEORY

I refer under that heading to the real vector field system

$$\begin{aligned}\mathcal{L}&=\underbrace{\tfrac{1}{4}(U^{\rho,\sigma}-U^{\sigma,\rho})(U_{\rho,\sigma}-U_{\sigma,\rho})}-\tfrac{1}{2}\varkappa^{2}U^{\sigma}U_{\sigma}\qquad(28.2\equiv 95.0)\\ &\qquad=\tfrac{1}{2}(U^{\rho,\sigma}U_{\rho,\sigma}-U^{\rho,\sigma}U_{\sigma,\rho})\end{aligned}$$

[55] See again (75).

into the theory of which the antisymmetric tensor field $G^{\mu\nu} \equiv \partial^\mu U^\nu - \partial^\nu U^\mu$ enters simply as a notational device. Introducing

$$\left.\begin{aligned} x^\mu &\longrightarrow x^\mu + \mathcal{X}^{\mu\alpha}\cdot\delta\omega_\alpha \quad \text{with} \quad \mathcal{X}^{\mu\alpha} = g^{\mu\alpha} \\ U^\mu &\longrightarrow U^\mu + \Phi^{\mu\alpha}\cdot\delta\omega_\alpha \quad \text{with} \quad \Phi^{\mu\alpha} = 0 \end{aligned}\right\} \tag{95.1}$$

into the generic

$$\begin{aligned} J^{\mu\alpha} =& \pi^{\mu\nu}\big\{\Phi_\nu{}^\alpha - U_{\nu,\sigma}\mathcal{X}^{\sigma\alpha}\big\} + \mathfrak{L}\mathcal{X}^{\mu\alpha} \\ & \pi^{\mu\nu} \equiv \partial\mathfrak{L}/\partial U_{\nu,\mu} \end{aligned} \tag{95.2}$$

we notice that $\pi^{\mu\nu} = U^{\nu,\mu} - U^{\mu,\nu} \equiv G^{\mu\nu}$ in the present instance, and that the Lagrangian can be notated $\mathfrak{L} = \frac{1}{4}\pi^{\rho\sigma}\pi_{\rho\sigma} - \frac{1}{2}\varkappa^2 U^\sigma U_\sigma$, and so obtain

$$\begin{aligned} \mathcal{T}^{\mu\alpha} &= \pi^\mu{}_\nu U^{\nu,\alpha} - \mathfrak{L}g^{\mu\alpha} \\ &\neq \mathcal{T}^{\alpha\mu} \end{aligned} \tag{95.3}$$

The infinitesimal Lorentz map reads

$$\left.\begin{aligned} x^\mu &\longrightarrow x^\mu + \tfrac{1}{2}\mathcal{X}^{\mu\alpha\beta}\cdot\delta\Omega_{\alpha\beta} \quad \text{with} \quad \mathcal{X}^{\mu\alpha\beta} = (g^{\mu\alpha}\,x^\beta - g^{\mu\beta}\,x^\alpha) \\ U^\mu &\longrightarrow U^\mu + \tfrac{1}{2}\Phi^{\mu\alpha\beta}\cdot\delta\Omega_{\alpha\beta} \quad \text{with} \quad \Phi^{\mu\alpha\beta} = (g^{\mu\alpha}U^\beta - g^{\mu\beta}U^\alpha) \end{aligned}\right\} \tag{95.4}$$

which when introduced into the generic

$$J^{\mu\alpha\beta} = \pi^{\mu\nu}\big\{\Phi_\nu{}^{\alpha\beta} - U_{\nu,\sigma}\mathcal{X}^{\sigma\alpha\beta}\big\} + \mathfrak{L}\mathcal{X}^{\mu\alpha\beta} \tag{95.5}$$

give

$$\begin{aligned} \mathcal{J}^{\mu\alpha\beta} =& \quad \tfrac{1}{\mathrm{c}}\Big[\big(-\pi^\mu{}_\nu U^{\nu,\alpha} + \mathfrak{L}g^{\mu\alpha}\big)x^\beta - \big(-\pi^\mu{}_\nu U^{\nu,\beta} + \mathfrak{L}g^{\mu\beta}\big)x^\alpha\Big] \\ & + \tfrac{1}{\mathrm{c}}\Big[\pi^{\mu\alpha}U^\beta - \pi^{\mu\beta}U^\alpha\Big] \\ =& \tfrac{1}{\mathrm{c}}\big(x^\alpha\mathcal{T}^{\mu\beta} - x^\beta\mathcal{T}^{\mu\alpha}\big) + S^{\mu\alpha\beta} \qquad (95.6) \\ & S^{\mu\alpha\beta} = \tfrac{1}{\mathrm{c}}\big(G^{\mu\alpha}U^\beta - G^{\mu\beta}U^\alpha\big) \qquad (95.7) \end{aligned}$$

We have now in hand all we need to accomplish the *symmetrization* of $\mathcal{T}^{\mu\alpha}$. The generic equation (83) becomes in the present instance

$$\begin{aligned} H^{\sigma\mu\nu} &= \tfrac{1}{2}c\big\{\mathcal{S}^{\sigma\mu\nu} + \mathcal{S}^{\mu\nu\sigma} - \mathcal{S}^{\nu\sigma\mu}\big\} \\ &= \tfrac{1}{2}\big\{G^{\sigma\mu}U^\nu - G^{\sigma\nu}U^\mu + G^{\mu\nu}U^\sigma - G^{\mu\sigma}U^\nu - G^{\nu\sigma}U^\mu + G^{\nu\mu}U^\sigma\big\} \\ &= G^{\sigma\mu}U^\nu \quad \text{after cancellations resulting from } G^{\mu\nu} = -G^{\nu\mu} \end{aligned} \tag{95.8}$$

Belinfante's (84) now gives

$$\begin{aligned} \mathcal{T}^{\mu\alpha} \longrightarrow T^{\mu\alpha} &= \mathcal{T}^{\mu\alpha} + \partial_\sigma H^{\sigma\mu\alpha} \\ & \partial_\sigma H^{\sigma\mu\alpha} = U^\alpha\partial_\sigma G^{\sigma\mu} + G^{\sigma\mu}\partial_\sigma U^\alpha \end{aligned} \tag{95.9}$$

Drawing upon the field equations and (once again) upon $G^{\sigma\mu} = -G^{\mu\sigma}$, we have

$$= -\varkappa^2 U^\alpha U^\mu - G^\mu{}_\sigma U^{\alpha,\sigma}$$

giving

$$= G^\mu{}_\sigma U^{\sigma,\alpha} - \mathfrak{L} g^{\mu\alpha} - \varkappa^2 U^\mu U^\alpha - G^\mu{}_\sigma U^{\alpha,\sigma}$$

whence

$$T^{\mu\alpha} = G^\mu{}_\sigma G^{\alpha\sigma} - \mathfrak{L} g^{\mu\alpha} - \varkappa^2 U^\mu U^\alpha \tag{95.10}$$

which is manifestly symmetric:

$$T^{\mu\alpha} = T^{\alpha\mu} \tag{95.11}$$

The implied energy density function is

$$\begin{aligned}
\mathcal{E} = T^{00} &= G^0{}_\sigma G^{0\sigma} - \left(\tfrac{1}{4} G^{\sigma\rho} G_{\sigma\rho} - \tfrac{1}{2}\varkappa^2 U^\sigma U_\sigma\right) - \varkappa^2 U^0 U^0 \\
&= \left(- G_{01}^2 - G_{02}^2 - G_{03}^2\right) \\
&\qquad + \tfrac{1}{2}\left(G_{01}^2 + G_{02}^2 + G_{03}^2 - G_{12}^2 - G_{23}^2 - G_{31}^2\right) \\
&\qquad + \tfrac{1}{2}\varkappa^2\left(U_0^2 - U_1^2 - U_2^2 - U_3^2\right) - \varkappa^2 U_0^2 \\
&= -\tfrac{1}{2}\Big[\left(G_{01}^2 + G_{02}^2 + G_{03}^2 + G_{12}^2 + G_{23}^2 + G_{31}^2\right) \\
&\qquad\qquad + \varkappa^2\left(U_0^2 + U_1^2 + U_2^2 + U_3^2\right)\Big] \leqslant 0
\end{aligned} \tag{95.12}$$

Evidently the

Procca theory requires a final sign-reversal to bring the regularization process to completion:

$$\begin{aligned}
T^{\mu\alpha} \longrightarrow \mathring{T}^{\mu\alpha} &\equiv -T^{\mu\alpha} \\
&= G^\mu{}_\sigma G^{\sigma\alpha} + \mathfrak{L} g^{\mu\alpha} + \varkappa^2 U^\mu U^\alpha
\end{aligned} \tag{95.13}$$

This step could be averted if we backed up to (95.0) and flipped the sign of the Lagrangian, but such a step might seem unmotivated; the moral is: *The sign of the Lagrangian makes a difference—not to the equations of motion, but to some of their deeper formal ramifications*. Note also that while $\mathfrak{T}^{00}$ and T^{00} assign identical values to the *total* energy, they differ by a term

$$\begin{aligned}
\partial_\sigma H^{\sigma 00} &= -\varkappa^2 U^0 U^0 - G^0{}_\sigma U^{0,\sigma} \\
&= -\varkappa^2 U_0^2 + \left(G_{01} U_{0,1} + G_{02} U_{0,2} + G_{03} U_{0,3}\right) \\
&= -\varkappa^2 U_0^2 - \left(U_{0,1}^2 + U_{0,2}^2 + U_{0,3}^2\right) + \left(U_{1,0} U_{0,1} + U_{2,0} U_{0,2} + U_{3,0} U_{0,3}\right) \\
&= -\varkappa^2 U_0^2 - \tfrac{1}{2}\left(U_{0,1}^2 + U_{0,2}^2 + U_{0,3}^2\right) - \tfrac{1}{2}\left(G_{01}^2 + G_{02}^2 + G_{03}^2\right) \\
&\qquad + \tfrac{1}{2}\left(U_{1,0}^2 + U_{2,0}^2 + U_{3,0}^2\right)
\end{aligned}$$

which is not only non-zero but *sign-indefinite*:[56] prior to regularization we stood at risk of violating the principle that energy density must be non-negative. Spin

[56] In the preceding argument I have used

$$G_{\mu\nu}^2 = (U_{\nu,\mu} - U_{\mu,\nu})^2 = U_{\mu,\nu}^2 + U_{\nu,\mu}^2 - 2U_{\mu,\nu} U_{\nu,\mu} \geqslant 0$$

structure was incorporated into the design of $\mathring{T}^{\mu\alpha}$, and is therefore incorporated also into the momental design of

$$\mathring{J}^{\mu\alpha\beta} \equiv \tfrac{1}{\mathrm{C}}(x^\alpha \mathring{T}^{\mu\beta} - x^\beta \mathring{T}^{\mu\alpha}) \tag{95.14}$$

Working from the generic equation (82) we are led back again to the spin tensor

$$\begin{aligned}\mathring{S}^{\mu\alpha\beta} = -S^{\mu\alpha\beta} &= \tfrac{1}{\mathrm{C}}\big(H^{\alpha\mu\beta} - H^{\beta\mu\alpha}\big) \\ &= \tfrac{1}{\mathrm{C}}\big(G^{\alpha\mu}U^\beta - G^{\beta\mu}U^\alpha\big)\end{aligned} \tag{95.15}$$

previously encountered at (95.7). The definition of the orbital angular momentum tensor is implicit in the relation

$$\mathring{J}^{\mu\alpha\beta} = \mathring{M}^{\mu\alpha\beta} + \mathring{S}^{\mu\alpha\beta} \tag{95.16}$$

$\partial_\mu \mathring{M}^{\mu\alpha\beta}$ and $\partial_\mu \mathring{S}^{\mu\alpha\beta}$ both fail to vanish, but in such a concerted way as to achieve $\partial_\mu \mathring{J}^{\mu\alpha\beta} = 0$.

ELECTRODYNAMICS

We look upon Maxwellian electrodynamics as "Procca theory in the zero-mass limit $\varkappa \downarrow 0$." By way of notational preparation, let us agree to write[57]

$$\begin{pmatrix} U^0 \\ U^1 \\ U^2 \\ U^3 \end{pmatrix} = \begin{pmatrix} \phi \\ \mathfrak{A} \end{pmatrix} \quad \text{and} \quad \|G^{\mu\nu}\| = \begin{pmatrix} 0 & -\mathfrak{E}_1 & -\mathfrak{E}_2 & -\mathfrak{E}_3 \\ \mathfrak{E}_1 & 0 & -\mathfrak{B}_3 & \mathfrak{B}_2 \\ \mathfrak{E}_2 & \mathfrak{B}_3 & 0 & -\mathfrak{B}_1 \\ \mathfrak{E}_3 & -\mathfrak{B}_2 & \mathfrak{B}_1 & 0 \end{pmatrix} \tag{96}$$

The Procca field equations (32)

$$\begin{gathered} G^{\mu\nu} = \partial^\mu U^\nu - \partial^\nu U^\mu \\ \partial^\lambda G^{\mu\nu} + \partial^\mu G^{\nu\lambda} + \partial^\nu G^{\lambda\mu} = 0 \\ \partial_\mu G^{\mu\nu} + \varkappa^2 U^\nu = 0 \\ \partial_\nu U^\nu = 0 \end{gathered}$$

can in this notation be expressed

$$\begin{gathered} \boldsymbol{\mathfrak{E}} = -\boldsymbol{\nabla}\phi - \tfrac{1}{\mathrm{C}}\tfrac{\partial}{\partial t}\boldsymbol{\mathfrak{A}} \qquad \text{and} \qquad \boldsymbol{\mathfrak{B}} = \boldsymbol{\nabla}\times\boldsymbol{\mathfrak{A}} \\ \boldsymbol{\nabla}\times\boldsymbol{\mathfrak{E}} + \tfrac{1}{\mathrm{C}}\tfrac{\partial}{\partial t}\boldsymbol{\mathfrak{B}} = \mathbf{0} \qquad \text{and} \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\mathfrak{B}} = 0 \\ \boldsymbol{\nabla}\cdot\boldsymbol{\mathfrak{E}} = -\varkappa^2\phi \qquad \text{and} \qquad \boldsymbol{\nabla}\times\boldsymbol{\mathfrak{B}} - \tfrac{1}{\mathrm{C}}\tfrac{\partial}{\partial t}\boldsymbol{\mathfrak{E}} = -\varkappa^2\boldsymbol{\mathfrak{A}} \\ \tfrac{1}{\mathrm{C}}\tfrac{\partial}{\partial t}\phi + \boldsymbol{\nabla}\cdot\boldsymbol{\mathfrak{A}} = 0 \end{gathered}$$

In the limit $\varkappa \downarrow 0$ these equations assume precisely the form of the *Maxwell's*

(continued from the preceding page) to obtain

$$\begin{aligned}\big(U_{1,0}U_{0,1} + U_{2,0}U_{0,2} + U_{3,0}U_{0,3}\big) &= \tfrac{1}{2}\big(U_{0,1}^2 + U_{0,2}^2 + U_{0,3}^2 + U_{1,0}^2 + U_{2,0}^2 + U_{3,0}^2\big) \\ &\quad - \tfrac{1}{2}\big(G_{01}^2 + G_{02}^2 + G_{03}^2\big) \\ &\leqslant \tfrac{1}{2}\big(U_{0,1}^2 + U_{0,2}^2 + U_{0,3}^2 + U_{1,0}^2 + U_{2,0}^2 + U_{3,0}^2\big)\end{aligned}$$

[57] Compare CLASSICAL ELECTRODYNAMICS (1980), pp. 162 & 373.

equations in the absence of sources,[58] joined by the Lorentz gauge condition as an automatic tag-along. We have

$$\begin{aligned} \mathring{T}^{00} &= \text{energy density} \\ &= \tfrac{1}{2}\big\{\mathfrak{E}^2 + \mathfrak{B}^2 + \varkappa^2\big(\phi^2 + \mathfrak{A}^2\big)\big\} \qquad (97.1) \\ \tfrac{1}{\mathrm{c}}\begin{pmatrix} \mathring{T}^{01} \\ \mathring{T}^{02} \\ \mathring{T}^{03} \end{pmatrix} &= \text{momentum density vector } \boldsymbol{P} \\ &= \tfrac{1}{\mathrm{c}}\big\{\boldsymbol{\mathfrak{E}}\times\boldsymbol{\mathfrak{B}} + \varkappa^2\phi\boldsymbol{\mathfrak{A}}\big\} \qquad (97.2) \end{aligned}$$

which again give back familiar results in the limit $\varkappa \downarrow 0$. Less familiar is the construction (see again (95.15))

$$\begin{aligned} \tfrac{1}{\mathrm{c}}\begin{pmatrix} \mathring{S}^{023} \\ \mathring{S}^{031} \\ \mathring{S}^{012} \end{pmatrix} = \tfrac{1}{\mathrm{c}}\begin{pmatrix} G^{20}U^3 - G^{30}U^2 \\ G^{30}U^1 - G^{10}U^3 \\ G^{10}U^2 - G^{20}U^1 \end{pmatrix} &= \text{spin density vector } \boldsymbol{S} \\ &= \tfrac{1}{\mathrm{c}}\boldsymbol{\mathfrak{E}}\times\boldsymbol{\mathfrak{A}} \qquad (98) \end{aligned}$$

This result is at least dimensionally correct,[59] but is nonetheless puzzling in this profound respect: (98) contains no allusion to the mass parameter $\varkappa$, and therefore remains unchanged in the electromagnetic limit, where we expect to obtain

$$\boldsymbol{S} = \tfrac{1}{\mathrm{c}}\boldsymbol{E}\times\boldsymbol{A} \qquad \textit{but only in the Lorentz gauge}$$

One would like

- to develop a *manifestly gauge-invariant* description of $\boldsymbol{S}$, else
- to identify *properties* of $\boldsymbol{S}$ which are gauge-invariant (therefore physical).

But those are assignments which I must save for another occasion.[60] It is a

[58] Curiously, if we strike the first line and make the notational/conceptual adjustments

$$-\varkappa^2\phi \longmapsto \rho \qquad \text{and} \qquad -\varkappa^2\boldsymbol{\mathfrak{A}} \longmapsto \tfrac{1}{\mathrm{c}}\mathbf{j}$$

we obtain

$$\begin{gathered} \boldsymbol{\nabla}\times\boldsymbol{\mathfrak{E}} + \tfrac{1}{\mathrm{c}}\tfrac{\partial}{\partial t}\boldsymbol{\mathfrak{B}} = \mathbf{0} \qquad \text{and} \qquad \boldsymbol{\nabla}\cdot\boldsymbol{\mathfrak{B}} = 0 \\ \boldsymbol{\nabla}\cdot\boldsymbol{\mathfrak{E}} = \rho \qquad \text{and} \qquad \boldsymbol{\nabla}\times\boldsymbol{\mathfrak{B}} - \tfrac{1}{\mathrm{c}}\tfrac{\partial}{\partial t}\boldsymbol{\mathfrak{E}} = \tfrac{1}{\mathrm{c}}\mathbf{j} \\ \tfrac{\partial}{\partial t}\rho + \boldsymbol{\nabla}\cdot\mathbf{j} = 0 \end{gathered}$$

which are precisely Maxwell's equations *in the presence* of sources. What was formerly the Lorentz gauge condition has become the charge continuity equation. The introduction of potentials ϕ and $\mathfrak{A}$ acquires now the status of an unexploited option.

[59] Working from $[G^2] = [\varkappa^2U^2] =$ (energy density) and $[\varkappa] =$ (length)$^{-1}$ we find $[\frac{1}{\mathrm{c}}GU] = [\frac{1}{\mathrm{c}}\mathfrak{E}\,\mathfrak{A}] =$ (energy $\cdot$ time)/(length)^3 =action density.

[60] In the meantime, see Soper,[46] p. 115; Corson,[46] p. 81; or Bjørn Felsager, *Geometry, Particles & Fields* (1997), §3.6.

lesson of experience that when $\boldsymbol{A}$-potentials step nakedly onto the stage we can expect odd goings-on, and that our careful attention will be rewarded.[61]

DIRAC FIELD

We revisit the system

$$\mathfrak{L} = -\hbar c\Big\{ i\tfrac{1}{2}\{\tilde{\psi}_{,\mu}\boldsymbol{\gamma}^{\mu}\psi - \tilde{\psi}\boldsymbol{\gamma}^{\mu}\psi_{,\mu}\} + \varkappa\,\tilde{\psi}\psi\Big\} \qquad (56 \equiv 99.0)$$

where ψ is now a 4-component spinor field. Introducing

$$\left.\begin{array}{lll} x^{\mu} \longrightarrow x^{\mu} + \mathcal{X}^{\mu\alpha}\cdot\delta\omega_{\alpha} & \text{with} & \mathcal{X}^{\mu\alpha} = g^{\mu\alpha} \\ \psi^{a} \longrightarrow \psi^{a} + \Phi^{a\alpha}\cdot\delta\omega_{\alpha} & \text{with} & \Phi^{a\alpha} = 0 \\ \tilde{\psi}^{a} \longrightarrow \tilde{\psi}^{a} + \tilde{\Phi}^{a\alpha}\cdot\delta\omega_{\alpha} & \text{with} & \tilde{\Phi}^{a\alpha} = 0 \end{array}\right\} \qquad (99.1)$$

into the generic

$$J^{\mu\alpha} = -\Big[\pi^{\mu}{}_{a}\{\Phi^{a\alpha} - \psi^{a}{}_{,\sigma}\mathcal{X}^{\sigma\alpha}\} + \{\tilde{\Phi}^{a\alpha} - \tilde{\psi}^{a}{}_{,\sigma}\mathcal{X}^{\sigma\alpha}\}\tilde{\pi}^{\mu}{}_{a} + \mathfrak{L}\mathcal{X}^{\mu\alpha}\Big] \qquad (99.2)$$
$$\pi^{\mu}{}_{a} \equiv \partial\mathfrak{L}/\partial\psi^{a}{}_{,\mu} \qquad\qquad \tilde{\pi}^{\mu}{}_{a} \equiv \partial\mathfrak{L}/\partial\tilde{\psi}^{a}{}_{,\mu}$$

we observe that

$$\pi^{\mu} = +\tfrac{1}{2}i\hbar c\tilde{\psi}\boldsymbol{\gamma}^{\mu} \qquad \text{and} \qquad \tilde{\pi}^{\mu} = -\tfrac{1}{2}i\hbar c\boldsymbol{\gamma}^{\mu}\psi$$

in the present instance, and that the Lagrangian can for the purposes at hand be notated

$$\begin{aligned} \mathfrak{L} &= \pi^{\sigma}\psi_{,\sigma} + \tilde{\psi}_{,\sigma}\tilde{\pi}^{\sigma} - mc^{2}\tilde{\psi}\psi \\ &= 0 \quad \text{numerically, as established at (58)} \end{aligned}$$

So (dropping the $\mathfrak{L}g^{\mu\alpha}$-term) we have

$$\begin{aligned} \mathcal{T}_{\mu\alpha} &= \pi_{\mu}\psi_{,\alpha} + \tilde{\psi}_{,\alpha}\tilde{\pi}_{\mu} \\ &= \tfrac{1}{2}i\hbar c\{\tilde{\psi}\boldsymbol{\gamma}_{\mu}\psi_{,\alpha} - \tilde{\psi}_{,\alpha}\boldsymbol{\gamma}_{\mu}\psi\} \\ &\neq \mathcal{T}_{\alpha\mu} \end{aligned} \qquad (99.3)$$

We are forced therefore to look to the spin structure of the Dirac field, as a first step toward Belinfante symmetrization (which in this special case Pauli[62] had accomplished by *ad hoc* methods already in 1933). The infinitesimal Lorentz map reads

$$\begin{array}{lll} x^{\mu} \longrightarrow x^{\mu} + \frac{1}{2}\mathcal{X}^{\mu\alpha\beta}\cdot\delta\Omega_{\alpha\beta} & \text{with} & \mathcal{X}^{\mu\alpha\beta} = (g^{\mu\alpha}\,x^{\beta} - g^{\mu\beta}\,x^{\alpha}) \\ \psi^{a} \longrightarrow \psi^{a} + \frac{1}{2}\Phi^{a\alpha\beta}\cdot\delta\Omega_{\alpha\beta} & \text{with} & \Phi^{a\alpha\beta} = B^{a}{}_{b}{}^{\alpha\beta}\psi^{b} \end{array}$$

[61] See, for example, the discussion of the Aharonov-Bohm effect in Griffiths' *Introduction to Quantum Mechanics* (1994), §10.2.4.

[62] See p. 235 in "Die allgemeinen Prinzipien der Wellenmechanik," *Handbuch der Physik* (2nd edition) **24/1** (1933), which was reprinted as a separate volume in 1950.

where preliminary information relating to the structure functions $\Phi^{a\alpha\beta}$ has been harvested from (75.2), but it remains to figure out the designs of the 4×4 matrices $\boldsymbol{B}^{\alpha\beta}\equiv\|B^a{}_b{}^{\alpha\beta}\|$; that, however, is easily accomplished, since the main work has already been done: returning with (74) to (72), we have

$$\begin{aligned}\boldsymbol{B}&=\tfrac{1}{8}A^{\mu\nu}(\boldsymbol{\gamma}_\mu\boldsymbol{\gamma}_\nu-\boldsymbol{\gamma}_\nu\boldsymbol{\gamma}_\mu)\\ A^{\mu\nu}&=\tfrac{1}{2}(g^{\mu\alpha}g^{\beta\nu}-g^{\mu\beta}g^{\alpha\nu})\delta\Omega_{\alpha\beta}\end{aligned}$$

giving

$$\begin{aligned}\boldsymbol{B}&=\tfrac{1}{2}\boldsymbol{B}^{\alpha\beta}\delta\Omega_{\alpha\beta}\\ \boldsymbol{B}^{\alpha\beta}&=\tfrac{1}{4}(\boldsymbol{\gamma}^\alpha\boldsymbol{\gamma}^\beta-\boldsymbol{\gamma}^\beta\boldsymbol{\gamma}^\alpha)\end{aligned}$$

From $\tilde\psi\equiv\psi^\dagger\boldsymbol{G}$ it follows that if $\delta_\Omega\psi=\boldsymbol{B}\psi$ then $\delta_\Omega\tilde\psi=\tilde\psi\tilde{\boldsymbol{B}}$, with $\tilde{\boldsymbol{B}}\equiv\boldsymbol{G}^{-1}\boldsymbol{B}^\dagger\boldsymbol{G}$. But the defining properties[63] of $\boldsymbol{G}$ entail $\tilde{\boldsymbol{\gamma}}^\mu=\boldsymbol{\gamma}^\mu$, from which it follows readily that $\tilde{\boldsymbol{B}}^{\alpha\beta}=\boldsymbol{B}^{\beta\alpha}=-\boldsymbol{B}^{\alpha\beta}$. The action of the infinitesimal Lorentz map can therefore be described

$$\left.\begin{aligned}x^\mu&\longrightarrow x^\mu+\tfrac{1}{2}\mathcal{X}^{\mu\alpha\beta}\cdot\delta\Omega_{\alpha\beta}\quad\text{with}\quad\mathcal{X}^{\mu\alpha\beta}=(g^{\mu\alpha}\,x^\beta-g^{\mu\beta}\,x^\alpha)\\ \psi&\longrightarrow\psi+\tfrac{1}{2}\Phi^{\alpha\beta}\cdot\delta\Omega_{\alpha\beta}\quad\text{with}\quad\Phi^{\alpha\beta}=+\tfrac{1}{4}(\boldsymbol{\gamma}^\alpha\boldsymbol{\gamma}^\beta-\boldsymbol{\gamma}^\beta\boldsymbol{\gamma}^\alpha)\psi\\ \tilde\psi&\longrightarrow\tilde\psi+\tfrac{1}{2}\tilde\Phi^{\alpha\beta}\cdot\delta\Omega_{\alpha\beta}\quad\text{with}\quad\tilde\Phi^{\alpha\beta}=-\tfrac{1}{4}\tilde\psi(\boldsymbol{\gamma}^\alpha\boldsymbol{\gamma}^\beta-\boldsymbol{\gamma}^\beta\boldsymbol{\gamma}^\alpha)\end{aligned}\right\}\tag{99.4}$$

We have achieved notational simplicity by surpressing the superscripts ^{a} that distinguish the components of the spinor field ψ, and would achieve more by introducing the $\alpha\beta$-antisymmeteric array of 4×4 matrices

$$\boldsymbol{\sigma}^{\alpha\beta}\equiv\tfrac{1}{2i}(\boldsymbol{\gamma}^\alpha\boldsymbol{\gamma}^\beta-\boldsymbol{\gamma}^\beta\boldsymbol{\gamma}^\alpha)$$

where the i has been introduced so as to achieve $\tilde{\boldsymbol{\sigma}}^{\alpha\beta}=\boldsymbol{\sigma}^{\alpha\beta}$. Returning with this information to the generic

$$J^{\mu\alpha\beta}=-\big\{\pi^\mu\psi_{,\sigma}+\tilde\psi_{,\sigma}\tilde\pi^\mu\big\}\mathcal{X}^{\sigma\alpha\beta}+\big\{\pi^\mu\Phi^{\alpha\beta}+\tilde\Phi^{\alpha\beta}\tilde\pi^\mu\big\}\tag{99.5}$$

we obtain

$$\mathcal{J}^{\mu\alpha\beta}=\tfrac{1}{\mathrm{C}}(x^\alpha\mathcal{T}^{\mu\beta}-x^\beta\mathcal{T}^{\mu\alpha})+S^{\mu\alpha\beta}\tag{99.6}$$

$$S^{\mu\alpha\beta}=-\tfrac{1}{4}\hbar\,\tilde\psi(\boldsymbol{\gamma}^\mu\boldsymbol{\sigma}^{\alpha\beta}+\boldsymbol{\sigma}^{\alpha\beta}\boldsymbol{\gamma}^\mu)\,\psi\tag{99.7}$$

With the Dirac spin tensor $S^{\mu\alpha\beta}$ now in hand,[64] we are in position to undertake the symmetrization of $\mathcal{T}^{\mu\alpha}$. The generic equation (83) becomes

$$\begin{aligned}H^{\sigma\mu\nu}&=-\tfrac{1}{8}\hbar c\tilde\psi\big\{\boldsymbol{\gamma}^\sigma\boldsymbol{\sigma}^{\mu\nu}+\boldsymbol{\sigma}^{\mu\nu}\boldsymbol{\gamma}^\sigma+\boldsymbol{\gamma}^\mu\boldsymbol{\sigma}^{\nu\sigma}+\boldsymbol{\sigma}^{\nu\sigma}\boldsymbol{\gamma}^\mu-\boldsymbol{\gamma}^\nu\boldsymbol{\sigma}^{\sigma\mu}-\boldsymbol{\sigma}^{\sigma\mu}\boldsymbol{\gamma}^\nu\big\}\psi\\ &=\tfrac{1}{8}i\hbar c\tilde\psi\big\{\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\sigma-\boldsymbol{\gamma}^\sigma\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\mu\big\}\psi\ \text{after algebraic simplifications}\\ &=-H^{\mu\sigma\nu}\quad\text{as required}\end{aligned}\tag{99.8}$$

---

[63] $\boldsymbol{G}^\dagger=\boldsymbol{G}$ and $(\boldsymbol{G}\boldsymbol{\gamma}^\mu)^\dagger=(\boldsymbol{G}\boldsymbol{\gamma}^\mu)$: see again the text preceding (55).

[64] For unaccountable reasons, (99.7) seems seldom to make an appearance in the standard literature, but see N. N. Bogoliubov & D. V. Shirkov, *Introduction to the Theory of Quantized Fields* (1959), eq. (7.31). It was from Chapter I of that volume that I learned my classical field theory, and the treatment of the classical Dirac field presented in §§6 & 7 still seems to me outstanding.

Drawing upon the fundamental anticommutation relation (53.1), one can show by quick calculation that

$$\{\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\sigma - \boldsymbol{\gamma}^\sigma\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\mu\} = \begin{cases} 2i\,\boldsymbol{\sigma}^{\mu\nu}\boldsymbol{\gamma}^\sigma + 2g^{\sigma\mu}\boldsymbol{\gamma}^\nu - 2g^{\sigma\nu}\boldsymbol{\gamma}^\mu \\ 2i\,\boldsymbol{\gamma}^\sigma\boldsymbol{\sigma}^{\mu\nu} - 2g^{\sigma\mu}\boldsymbol{\gamma}^\nu + 2g^{\sigma\nu}\boldsymbol{\gamma}^\mu \end{cases}$$

So we have

$$\begin{aligned}\vartheta_{\mu\nu} \equiv \partial_\sigma H^\sigma{}_{\mu\nu} = \quad & \tfrac{1}{4}i\hbar c\tilde{\psi}\big\{i\,\boldsymbol{\sigma}_{\mu\nu}\boldsymbol{\gamma}^\sigma\psi_{,\sigma} + \boldsymbol{\gamma}_\nu\psi_{,\mu} - \boldsymbol{\gamma}_\mu\psi_{,\nu}\big\} \\ & + \tfrac{1}{4}i\hbar c\big\{i\,\tilde{\psi}_{,\sigma}\boldsymbol{\gamma}^\sigma\boldsymbol{\sigma}_{\mu\nu} - \tilde{\psi}_{,\mu}\boldsymbol{\gamma}_\nu + \tilde{\psi}_{,\nu}\boldsymbol{\gamma}_\mu\big\}\psi\end{aligned}$$

The field equations $\boldsymbol{\gamma}^\sigma\psi_{,\sigma} = -i\varkappa\psi$ and $\tilde{\psi}_{,\sigma}\boldsymbol{\gamma}^\sigma = +i\varkappa\tilde{\psi}$ are invoked to bring about cancellation of the $\boldsymbol{\sigma}_{\mu\nu}$-factors, leaving

$$\vartheta_{\mu\nu} = -\tfrac{1}{4}i\hbar c\big\{(\tilde{\psi}\boldsymbol{\gamma}_\mu\psi_{,\nu} - \tilde{\psi}_{,\nu}\boldsymbol{\gamma}_\mu\psi) - (\tilde{\psi}\boldsymbol{\gamma}_\nu\psi_{,\mu} - \tilde{\psi}_{,\mu}\boldsymbol{\gamma}_\nu\psi)\big\} \tag{99.9}$$

which is manifestly real and (in this instance) manifestly $\mu\nu$-antisymmetric:

$$\vartheta_{\mu\nu} = \vartheta^*_{\mu\nu} \quad \text{and} \quad \vartheta_{\mu\nu} = -\vartheta_{\nu\mu}$$

We now resolve the stress-energy tensor $\mathcal{T}_{\mu\nu}$, as described at (99.3), into its symmetric and antisymmetric parts

$$\mathcal{T}_{\mu\nu} = \tfrac{1}{2}\big(\mathcal{T}_{\mu\nu} + \mathcal{T}_{\nu\mu}\big) + \tfrac{1}{2}\big(\mathcal{T}_{\mu\nu} - \mathcal{T}_{\nu\mu}\big)$$

and notice that $\frac{1}{2}\big(\mathcal{T}_{\mu\nu} - \mathcal{T}_{\nu\mu}\big)$ is precisely the negative of $\vartheta_{\mu\nu}$! Therefore Belinfante's

$$\begin{aligned} T_{\mu\nu} &= \mathcal{T}_{\mu\nu} + \vartheta_{\mu\nu} \\ &= \tfrac{1}{2}\big(\mathcal{T}_{\mu\nu} + \mathcal{T}_{\nu\mu}\big) \quad \text{is assuredly symmetric, and given by} \\ &= \tfrac{1}{4}i\hbar c\Big\{(\tilde{\psi}\boldsymbol{\gamma}_\mu\psi_{,\nu} - \tilde{\psi}_{,\nu}\boldsymbol{\gamma}_\mu\psi) + (\tilde{\psi}\boldsymbol{\gamma}_\nu\psi_{,\mu} - \tilde{\psi}_{,\mu}\boldsymbol{\gamma}_\nu\psi)\Big\} \end{aligned} \tag{99.10}$$

In short: Belinfanate's procedure works... but one must *work* to make it work!

The energy density of a Dirac field, by this accounting, becomes

$$\mathcal{E} = \tfrac{1}{2}i\hbar c\,(\tilde{\psi}\boldsymbol{\gamma}_0\psi_{,0} - \tilde{\psi}_{,0}\boldsymbol{\gamma}_0\psi) \tag{99.12}$$

This is precisely the formula (57) we obtained *prior* to symmetrization:

$$T_{00} = \mathcal{T}_{00}$$

as implied by the atypical antisymmetry of the $\vartheta_{\mu\nu}$ to which the Dirac theory has just lead us. The occurance of mixed signs in (99.12) casts into doubt any hope that the Dirac theory might conform to the "principle of energy

non-negativity," and indeed, it stands in celebrated violation of that principle. This deep fact follows most readily from (99.12), according to which

$$\text{time-reversal sends } \mathcal{E} \longrightarrow -\mathcal{E}$$

and can be attributed to the circumstance that if $(E/c)^2 - p^2 = (mc)^2$ then

$$E = \pm c\sqrt{p^2 + (mc)^2}$$

It was to circumvent this seeming defect of the theory that Dirac (who imagined himself to be inventing a theory of electrons) invented the "electron sea," and invoked the exclusion principle to fill up all the negative energy states. Proper resolution of the problem had to await development of a theory of *quantized* Dirac fields; in that expanded setting it becomes possible to "reinterpret the problem away" (i.e., to assign the "negative energy states" to "antiparticles").

We look finally to the spin density of the classical Dirac field; i.e., to the expressions

$$\begin{aligned} R_1 &\equiv S^{001} & S_1 &\equiv S^{023} \\ R_2 &\equiv S^{002} & S_2 &\equiv S^{031} \\ R_3 &\equiv S^{003} & S_3 &\equiv S^{012} \end{aligned}$$

Look first to the former: (99.7) supplies $R_1 = \frac{1}{4}\hbar\,\tilde{\psi}(\boldsymbol{\gamma}^0\boldsymbol{\sigma}^{01} + \boldsymbol{\sigma}^{01}\boldsymbol{\gamma}^0)\psi$, but

$$\boldsymbol{\gamma}^0\boldsymbol{\sigma}^{01} + \boldsymbol{\sigma}^{01}\boldsymbol{\gamma}^0 = \boldsymbol{\gamma}^0\boldsymbol{\gamma}^0\boldsymbol{\gamma}^1 - \boldsymbol{\gamma}^0\boldsymbol{\gamma}^1\boldsymbol{\gamma}^0 + \boldsymbol{\gamma}^0\boldsymbol{\gamma}^1\boldsymbol{\gamma}^0 - \boldsymbol{\gamma}^1\boldsymbol{\gamma}^0\boldsymbol{\gamma}^0 = \mathbf{0}$$

Evidently

$$R_1 = R_2 = R_3 = 0$$

This is a remarkable fact, a fact upon which all inertial observers agree. . . but which they do not much talk about in the literature; it can be phrased this way: *the spin-analogs of* $\{K_1, K_2, K_3\}$*—which in discussion subsequent to (77) were found to refer to the center of mass motion of a field system—are for Dirac fields trivial*. Drawing upon (99.7), we have

$$\begin{aligned} S_1 &= \psi^\dagger \boldsymbol{\Sigma}_1 \psi \quad \text{with} \quad \boldsymbol{\Sigma}_1 \equiv -\tfrac{1}{4}\hbar\boldsymbol{G}(\boldsymbol{\gamma}^0\boldsymbol{\sigma}^{23} + \boldsymbol{\sigma}^{23}\boldsymbol{\gamma}^0) \\ S_2 &= \psi^\dagger \boldsymbol{\Sigma}_2 \psi \quad \text{with} \quad \boldsymbol{\Sigma}_2 \equiv -\tfrac{1}{4}\hbar\boldsymbol{G}(\boldsymbol{\gamma}^0\boldsymbol{\sigma}^{31} + \boldsymbol{\sigma}^{31}\boldsymbol{\gamma}^0) \\ S_3 &= \psi^\dagger \boldsymbol{\Sigma}_3 \psi \quad \text{with} \quad \boldsymbol{\Sigma}_3 \equiv -\tfrac{1}{4}\hbar\boldsymbol{G}(\boldsymbol{\gamma}^0\boldsymbol{\sigma}^{12} + \boldsymbol{\sigma}^{12}\boldsymbol{\gamma}^0) \end{aligned}$$

Recalling from (54) how the $\boldsymbol{\gamma}^\mu$ were defined, and from (55) how $\boldsymbol{G}$ was defined; recalling also the definitions $\boldsymbol{\sigma}^{\mu\mu}$ which were motivated by (99.4), and relying upon *Mathematica* to perform the matrix algebra, we obtain

$$\boldsymbol{\Sigma}_1 = \tfrac{1}{2}\hbar\begin{pmatrix} \boldsymbol{\sigma}_1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}_1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \tfrac{1}{2}\hbar\begin{pmatrix} \boldsymbol{\sigma}_2 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}_2 \end{pmatrix}, \quad \boldsymbol{\Sigma}_3 = \tfrac{1}{2}\hbar\begin{pmatrix} \boldsymbol{\sigma}_3 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\sigma}_3 \end{pmatrix}$$

The simplicity of this result is striking, and so are its implications, for it is immediately evident that $\boldsymbol{\Sigma}$-algebra duplicates the Pauli algebra; specifically

$$\begin{aligned}\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2 - \boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_1 &= i\hbar\,\boldsymbol{\Sigma}_3\\ \boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_3 - \boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_2 &= i\hbar\,\boldsymbol{\Sigma}_1\\ \boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_1 - \boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_3 &= i\hbar\,\boldsymbol{\Sigma}_2\end{aligned}$$

which inform us we are dealing with "angular momentum algebra." Moreover

$$\boldsymbol{\Sigma}_1^2 + \boldsymbol{\Sigma}_2^2 + \boldsymbol{\Sigma}_3^2 = \ell(\ell+1)\boldsymbol{I} \quad \text{with} \quad \ell = \tfrac{1}{2}$$

We have touched here upon ideas standard to elementary quantum mechanics,[65] and it is true, moreover, that the expressions

$$\iiint \psi^\dagger \boldsymbol{\Sigma}_1 \psi \, dx^1 dx^2 dx^3, \text{ etc.}$$

which describe the *total* spin of the Dirac field do resemble quantum mechanical expectation values. It's all the more important, therefore, to appreciate that we are at the moment doing classical field theory—*not* quantum mechanics—and that the $\hbar$'s which appear in our equations are simply abbreviations for a certain constellation

$$\hbar \equiv mc/\varkappa$$

of the dimensioned constants m, c and $\varkappa$ which enter as essential players into the classical theory of Dirac fields; $\hbar$, for present purposes, has *not* the status of a God-given constant of Nature, but might have any value.

Concluding remarks. All linear relativistic field theories assume a distinctly Dirac-like appearance (36) when cast into canonical form. One expects the discussion just concluded to provide, therefore, a pattern adaptable to the generality of such theories. Were one to undertake such a program, one would want in particular to identify the algebraic circumstances which the cases $\mathcal{E} \geqslant 0$ from the cases $\mathcal{E} \gtrless 0$.

In an expanded review of relativistic field theories one would expect to encounter some discussion of "zero-mass theories." We have seen that the limit $\varkappa \downarrow 0$ must be approached circumspectly, that massless fields display internal degrees of freedom (think of gauge freedom in the Maxwellian case) which are absent if $\varkappa \neq 0$. The foundations of the theory of massless fields were laid down

[65] See, for example, D. Griffiths, *Introduction to Quantum Mechanics* (1995), §§4.3 & 4.4.

by Hermann Weyl[66] and, in an alternative form, by Ettore Majorana.[67] Brief accounts of the subject can be found in several of the standard sources.[68]

It is entirely possible to retain Lorentz covariance but *abandon the linearity assumption* which has dominated the preceding discussion. The resulting theory is much more difficult, but for a period during the 1950's Heisenberg imagined it to be the shape of the future.[69] Heisenberg looked to theories of (roughly) the form $\boldsymbol{\gamma}^\mu \partial_\mu \psi + i\varkappa(\tilde{\psi}\boldsymbol{\gamma}_\mu\psi)\boldsymbol{\gamma}^\mu\psi = 0$. Fairly typical of the large (but largely inconsequential) literature is a paper[70] by G. Rosen, who has, for his own reasons, discussed the system $\mathfrak{L} = \frac{1}{2}g^{\alpha\beta}\varphi_{,\alpha}\varphi_{,\beta} + g\varphi^6$. Einstein's theory of gravitation provides the non-linear field theory *par excellence*, but there one abandons not only linearity but also special relativity. The theory of solitons presents occasional instances of non-linear relativistic fields, though that theory is for the most part studied non-relativistically, on spaces of reduced dimension.

---

[66] "Elektron und Gravitation," Zeits. für Phys. **56**, 330 (1929). It was entirely characteristic of Weyl to draw his motivation from general relativity, rather than from the still embryonic physics of elementary particles; recall that Pauli's "neutrino hypothesis" (1933) lay still four years in the future. Pauli criticized Weyl's "two component theory" on grounds that it was not parity symmetric, and it was for that precise reason that the theory—in the guise now of a theory of neutrinos—was revived by T. D. Lee & C. N. Yang: "Question of parity conservation in weak interactions," Phys. Rev. **104**, 254 (1956). It was, by the way, in that same classic paper that Weyl lay the foundation for what was to become gauge field theory; see Chapter 5 in L. O'Raifeartaigh, *The Dawning of Gauge Theory* (1997).

[67] "Teoria relativistica di particelle con momento intrinseco arbitrario," Nuovo Cimento **9**, 335 (1932); "Teoria simmetrica dell' elettrone e delpositrone," Nuovo Cimento **14**, 171 (1937). Majorana (1906–1938?) was a brilliant but habitually morose member of the stellar group which condensed around Fermi at the University of Rome in the late 1920's and early 1930's. The mystery of his disappearance, at age 32, has never been solved. Majorana's first paper is "pre-neutrino," but by the time of his second paper he had reason to be well-acquainted with that subject: Fermi had given the neutrino its name, and made it the central player in his "Versuch einer Theorie der β-Strahlen," Zeits. für Phys. **88**, 161 (1934.)

[68] See Schweber,[7] Chapter 5; Corson,[46] §29. The subject is treated also in Section III of W. Pauli & M. Fierz, "On the relativistic wave equations for particles of arbitrary spin in an electromagnetic field," Proc. Roy. Soc. **173A**, 211 (1939) and in G. Uhlenbeck & O. Laport, "Application of spinor analysis to the Maxwell & Dirac equations," Phys. Rev. **37**, 1380 (1931), both of which are classic.

[69] See his *Introduction to the Unified Field Theory of Elementary Particles* (1966). The theory advocated by Heisenberg was for a while embraced—but then publicly denounced—by Pauli.

[70] "Equations of motion in classical non-linear field theories," J. Math. Phys. **8**, 573 (1967).

Note should be made also of the "non-linear electrodynamics" which Max Born and Leopold Infeld were motivated to devevlop in the early 1930's.[71]

Relativistic linear field theory has been cultivated not as mathematical recreation, but because the subject—even prior to quantization—has important physical work to do. One has interest, therefore, not only in the formal design of such theories (the topic which has concerned us) but also in the *solutions* of the resulting field equations. The elaborate technology developed in response to the latter interest was first sketched by Jordan & Pauli,[72] and brought to its modern form by Schwinger.[73] The central objects are certain "invariant functions"—by nature Green's functions, made available by the linearity assumption.[74] The invariant functions developed in association with the classical theory see service also in the quantum theory of fields.

Information relating to each of the topics mentioned above can, as I have indicated, be found in the literature. I conclude with mention of a topic for which I am, on the other hand, able to cite no reference. Belinfante's procedure, as we have seen, acheives $\mathcal{T}^{\mu\nu}$ symmetrization by appeal to the spin structure of the field system. I interpret this to mean that it was the presence of "spin structure" which *served initially to destroy stress-energy symmetry*. Can one, in physical language, say anything illuminating about the *mechanism* by which spin disrupts symmetry? I pose that question as unfinished business, and think that a sharp answer (in—say—the Fermi/Weisskopf tradition) would serve very usefully to deepen our intuitive understanding of the dynamics of fields.[75]

---

[71] Their "Foundations of a new field theory" (Proc. Roy. Soc. **144A**, 425 (1934)) provides good exercise in the methods of classical field theory. Subsequent papers (Proc. Roy. Soc. **147A**, 522 (1934) and **150A**, 141 (1935)) discuss quantization of the theory; they lead to operators descriptive of the center of mass of the "new field" which fail to commute (as position operators are supposed to do). M. H. L. Pryce, in a companion paper ("Commuting coordinates in the new field theory," Proc. Roy. Soc. **150A**, 166 (1934)) shows that commutivity can be restored if one adopts modified definitions which entail addition of a term which refers to the spin of the field. Pryce's construction is in some ways anticipatory of Belinfante's. The Born-Leopold theory is discussed in its historical context by A. Sommerfeld in §37 of his *Electrodynamics* (1952).

[72] P. Jordan & W. Pauli,"Zur Quantenelektrodynamik ladungsfreier Felder," Zeits. für Phys. **47**, 151 (1928).

[73] J. Schwinger, "Quantum electrodynamics. II. Vacuum polarization & self-energy," Phys. Rev. **75**, 651 (1949). The material to which I refer is found in the Appendix.

[74] For surveys of the construction and properties of those functions, see RELATIVISTIC CLASSICAL FIELDS (1973), pp. 156–216; ANALYTICAL METHODS OF PHYSICS (1981), pp. 366–433 or any good quantum fields text; I particularly recommend Appendix A1 in J. M. Jauch & F. Rohrlich, *The Theory of Photons and Electrons* (1955).

[75] H. C. Ohanian's "What is spin?" (AJP **54**, 500 (1986)) takes a first step in the right direction, and is of independent interest.

3

CLASSICAL GAUGE FIELDS

Introduction. The theory of "gauge fields" (sometimes called "compensating fields"[1]) is today universally recognized to constitute one of the supporting pillars of fundamental physics, but it came into the world not with a revolutionary bang but with a sickly whimper, and took a long time to find suitable employment. It sprang from the brow of the youthful Hermann Weyl (1885–1955), who is generally thought of as a mathematician, but for the seminal importance of his contributions to general relativity and quantum mechanics—and, more generally, to the "geometrization of physics"—must be counted among the greatest physicists of the 20$^{\text{th}}$ Century. Weyl's initial motivation (1918) was to loosen up the mathematical apparatus of general relativity[2] just enough to find a natural dwelling place for electromagnetism. In 1927 Fritz London suggested that Weyl's idea rested more naturally upon quantum mechanics (then fresh out of the egg!) than upon general relativity, and in 1929 Weyl published a revised elaboration of his original paper—the classic "Elektron und Gravitation" to which I have already referred.[3] The influential Wolfgang Pauli became an ardent champion of the ideas put forward by Weyl, and it was *via* Pauli (whose "Wellenmechanik" article in the *Handbuch der Physik* (1933) had made a profound impression upon him) that those ideas

[1] See Section 21 in F. A. Kaempffer's charmingly eccentric *Concepts in Quantum Mechanics* (1965).

[2] Recall that Einstein's theory of gravitation had been completed only in 1915, and that its first observationial support was not forthcoming until 1919.

[3] The text, in English translation, can be found (together with historical commentary) in Lochlainn O'Raufeartaugh's splendid *The Dawning of Gauge Theory* (1997), which should be consulted for a much more balanced account of events than I can present here.

came to the attention of C. N. Yang, in the early 1950's. The attempt by Yang & Mills (1954) to construct a "gauge theory of nuclear forces" failed, for reasons (it became clear in retrospect) having to do with the fact that the nuclear force is too densely phenomenological—too far removed from fundamentals—to admit of any elegantly simple theory. The Yang–Mills theory did serve to bring gauge theory to the general attention of theorists, but several developments had to transpire...

- attention had to shift from the interaction of nucleons to the physics *interior* to nucleons (this development hinged upon the invention of the quark, by Gell-Mann and Zweig in 1964)
- the ideas had to come into place which made possible the development (by Weinberg and Salam in 1967) of a unified theory of electromagnetic and weak interactions[4]

...before it became evident (by the early 1970's) how gauge field theory fit within the Grand Scheme of Things.

The developments to which I have alluded, insofar as they refer to particle physics, are profoundly quantum mechanical. But the associated gauge field theory is, to a remarkable degree, susceptible to description in the language of *classical* field theory, and it is to that language—to the physics of "classical gauge fields"—that I here confine myself; $\hbar$'s will intrude, but they will always be "soft $\hbar$'s," inserted for dimensional reasons but stripped of their quantum mechanical burden.

Basic objective of the theory, as standardly conceived. It is a familiar fact that the physical output of quantum theory is phase insensitive—invariant, that is to say, under

$$\psi \longrightarrow \psi' \equiv e^{i\omega}\psi \tag{1}$$

We may attribute this circumstance to the reality of the Schrödinger Lagrangian

$$\mathcal{L} = \tfrac{1}{2}i\hbar(\psi_t^*\psi - \psi^*\psi_t) + \tfrac{\hbar^2}{2m}\boldsymbol{\nabla}\psi^*\cdot\boldsymbol{\nabla}\psi + \psi^*U\psi \tag{2}$$

from which at (1–91) we extracted the Noetherean conservation law

$$\tfrac{\partial}{\partial t}(\psi^*\psi) + \boldsymbol{\nabla}\cdot(\text{probability current}) = 0$$

If we adopt the polar representation $\psi = R\exp\left\{\frac{i}{\hbar}S\right\}$ then becomes[5]

$$\mathcal{L} = R^2\left[S_t + \tfrac{1}{2m}\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S + U\right] + \tfrac{\hbar^2}{2m}\boldsymbol{\nabla}R\cdot\boldsymbol{\nabla}R$$

which is manifestly invariant under this reformulation of (1):

$$S \longrightarrow S + \text{constant} \tag{3}$$

[4] For a good brief account of the developments to which I refer, see Chapter 1 in David Griffiths' *Introduction to Elementary Particles* (1987).

[5] See again (1–78).

Dropping the final term in the preceding Lagrangian, we obtain precisely the Lagrangian

$$\mathcal{L} = R^2 \left[S_t + \tfrac{1}{2m} \boldsymbol{\nabla} S \cdot \boldsymbol{\nabla} S + U \right] \tag{4}$$

which was seen at (1–108) to give rise to classical Hamilton-Jacobi theory; the evident invariance of that theory under (3) can by

$$S = \int L \, dt$$

be attributed to the well-known fact that the physical output from Lagrangian mechanics is insensitive to gauge transformations

$$L \longrightarrow L + \tfrac{d}{dt}(\text{anything})$$

Pauli, in a paper[6] which was influential in the history of thus subject, called

- physics-preserving transformations-by-*multiplicative*-adjustment (such as our (1)) "gauge transformations of the 1st type," and
- physics-preserving transformations-by-*additive*-adjustment (such as (3)) "gauge transformations of the 2nd type."

In gauge field theory the two tend to be joined at the hip, and in casual usage the term "gauge transformation" may refer to either.

Write (1) more carefully

$$\psi(x,t) \longrightarrow \psi'(x,t) \equiv e^{i\omega} \psi(x,t) \tag{5}$$

to emphasize the presumed $\{x,t\}$-independence of the phase factor, which we have in mind when we refer to the "global" character of the gauge transformation. The "local" analog of (5) reads

$$\psi(x,t) \longrightarrow \psi'(x,t) \equiv e^{ig\Omega(x,t)} \psi(x,t) \tag{6}$$

—the point being that *the phase factor is allowed now to vary from point to point*. Evidently $\psi^* \psi$ is invariant under (6), but from

$$\partial\psi(x,t) \longrightarrow \partial\psi'(x,t) \equiv e^{ig\Omega(x,t)} \big\{ \partial\psi(x,t) + ig[\partial\Omega(x,t)]\,\psi \big\} \tag{7}$$

we see that the adjustment (5)⟶(6) serves to disrupt the invariance of expressions assembled from derivatives (unless $\partial\Omega = 0$, which would take us back to the global theory). Gauge field theory presents a general mechanism for restoring gauge invariance to theories which the adjustment

$$\text{global} \longrightarrow \text{local}$$

[6] "Relativistic theories of elementary particles," Rev. Mod. Phys. **13**, 203 (1941). See the text subsequent to equations (23) in Part I, Section 2.

has served to disrupt. That mechanism, in its most frequently encountered (but, as will emerge, not its simplest) manifestation, can be described as follows:[7]

STEP ONE Make everywhere the substitutional replacement

$$\begin{array}{c}\partial_\mu \\ \downarrow \\ \partial_\mu - igA_\mu\end{array}$$

where $A_\mu(x)$ is a "gauge field" ("compensating field"), endowed with properties soon to be specified. Consider (7) to have, in consequence, become

$$(\partial_\mu - igA_\mu)\,\psi \longrightarrow (\partial_\mu - igA'_\mu)\,\psi' = e^{ig\Omega}\big\{(\partial_\mu - igA'_\mu)\,\psi + ig\tfrac{\partial\Omega}{\partial x^\mu}\psi\big\}$$

and STEP TWO assign to the "local gauge transformation" concept this enlarged meaning

$$\left.\begin{array}{l}\psi \;\longrightarrow\; \psi' = e^{ig\Omega}\cdot\psi \\ A_\mu \longrightarrow A'_\mu = A_\mu + \frac{\partial\Omega}{\partial x^\mu}\end{array}\right\} \tag{8}$$

so as to achieve

$$\mathcal{D}_\mu\psi \longrightarrow \mathcal{D}'_\mu\psi' = e^{ig\Omega}\cdot\mathcal{D}_\mu\psi \tag{9}$$

$$\mathcal{D}_\mu \equiv \partial_\mu - igA_\mu \tag{10}$$

which mimics the structure of the first of equations (8).

Given interest in a system $\mathcal{L}_0(\varphi,\partial\varphi)$, STEP THREE look to the modified system

$$\mathcal{L}_1(\varphi,\partial\varphi,A) \equiv \mathcal{L}_0(\varphi,\mathcal{D}\,\varphi) \tag{11}$$

which will be locally gauge invariant if the initial system was globally so.

To see how this works in a particular case, let us look to the relativistic complex scalar field system (2–19)

$$\mathcal{L}_0(\psi,\psi^*,\partial\psi,\partial\psi^*) = \tfrac{\hbar^2}{2m}\big\{g^{\alpha\beta}\psi^*_{,\alpha}\psi_{,\beta} - \varkappa^2\psi^*\psi\big\} \tag{12}$$

where I have set $K = mc^2/\varkappa^2 = \hbar^2/2m$ in order to achieve

$$[\psi^*\psi] = 1/(\text{length})^3$$

The invariance of L_0 under the *global* gauge transformation (1) is manifest, and was shown at (2–21) to entail conservation of the real-valued Noetherean current

$$Q^\mu \equiv \tfrac{\hbar}{m}g^{\mu\alpha}\left\{\frac{\psi^*_{,\alpha}\psi - \psi^*\psi_{,\alpha}}{2i}\right\} \tag{13}$$

[7] Gauge field theory is not intrinsically/essentially relativistic, but it is (like field theory generally) "relativistically predisposed," and I find that it serves expository simplicity to make use here of the notational conventions of relativity.

where the $\hbar/m$ was introduced in order to achieve $[Q] = 1/(\text{area}\cdot\text{time})$. To achieve *local* phase invariance we look to the modified system

$$\begin{aligned}\mathcal{L}_1(\psi,\psi^*,\partial\psi,\partial\psi^*,A) &= \tfrac{\hbar^2}{2m}\big\{g^{\alpha\beta}(\psi^*_{,\alpha}+igA_\alpha\psi^*)(\psi_{,\beta}-igA_\beta\psi)-\varkappa^2\psi^*\psi\big\}\\ &= \mathcal{L}_0(\psi,\psi^*,\partial\psi,\partial\psi^*)+\underbrace{ig\tfrac{\hbar^2}{2m}(\psi^*\psi_{,\alpha}-\psi^*_{,\alpha}\psi)A^\alpha}_{g\hbar Q_\alpha A^\alpha}+g^2\tfrac{\hbar^2}{2m}(\psi^*\psi)A_\alpha A^\alpha\end{aligned}$$

Looking to the equations of motion, we find by calculation that

$$\Big\{\partial_\nu\frac{\partial}{\partial\psi^*_{,\nu}}-\frac{\partial}{\partial\psi^*}\Big\}\mathcal{L}=0 \quad \text{becomes} \quad (g^{\alpha\beta}\mathcal{D}_\alpha\mathcal{D}_\beta+\varkappa^2)\psi=0 \tag{14.1}$$

$$\Big\{\partial_\nu\frac{\partial}{\partial\psi_{,\nu}}-\frac{\partial}{\partial\psi}\Big\}\mathcal{L}=0 \quad \text{gives the conjugated equation} \tag{14.2}$$

Finally—in what is perhaps the most amazingly productive step in the entire procedure—we STEP FOUR *launch the gauge field into motion* by introducing some

- quadratic
- gauge-invariant, and (in relativistic field theory also)
- Lorentz-invariant

∂A_μ-dependence into the Lagrangian. To that end, we note that

$$F_{\mu\nu}\equiv\partial_\mu A_\nu-\partial_\nu A_\mu \quad \text{is transparently gauge-invariant} \tag{15}$$

and that $F_{\alpha\beta}F^{\alpha\beta}$ answers to our other requirements; we look, therefore, to the twice-modified system

$$\begin{aligned}\mathcal{L}_2(\psi,\psi^*,\partial\psi,\partial\psi^*,A,\partial A) &= \mathcal{L}_0(\psi,\psi^*,\partial\psi,\partial\psi^*) \\ &\quad+\mathcal{L}_{\text{int}}(\psi,\psi^*,\partial\psi,\partial\psi^*,A)+\mathcal{L}(A,\partial A)\end{aligned} \tag{16.0}$$

where the "interaction term"

$$\mathcal{L}_{\text{int}}(\psi,\psi^*,\partial\psi,\partial\psi^*,A)=ig\tfrac{\hbar^2}{2m}(\psi^*\psi_{,\alpha}-\psi^*_{,\alpha}\psi)A^\alpha+g^2\tfrac{\hbar^2}{2m}(\psi^*\psi)A_\alpha A^\alpha \tag{16.1}$$

was developed already at the top of the page, and where the "free gauge field" will be governed by

$$\begin{aligned}\mathcal{L}(A,\partial A) &= \tfrac14 F_{\alpha\beta}F^{\alpha\beta}\\ &= \tfrac14 g^{\alpha\rho}g^{\beta\sigma}(A_{\beta,\alpha}-A_{\alpha,\beta})(A_{\sigma,\rho}-A_{\rho,\sigma})\\ &= \tfrac12(g^{\alpha\rho}g^{\beta\sigma}-g^{\alpha\sigma}g^{\beta\rho})A_{\alpha,\beta}A_{\rho,\sigma}\end{aligned} \tag{16.2}$$

Notice also that $\mathcal{L}(A,\partial A)$ is in fact A_μ-*independent*; we have been forced to omit an anticipated $(\varkappa^2A_\alpha A^\alpha)$-term for the simple but deeply consequential reason that

$$A_\alpha A^\alpha \quad \text{is not gauge-invariant}$$

Notice also that (16.2) requires $[F_{\alpha\beta}F^{\alpha\beta}] = $ (energy density), which entails $[A_\mu] = \sqrt{\text{energy/length}}$. And this—if we are to achieve $[gA_\mu] = 1/(\text{length})$, as required by the definition of $\mathcal{D}_\mu$—in turn entails

$$[g^2] = 1/(\text{energy}\cdot\text{length}) = [1/\hbar c] = 1/(\text{electric charge})^2$$

giving $[g\hbar c] = $ (electric charge). It is on this basis that we will adopt $e \equiv g\hbar c$ as a suggestive notational device.

Working now from the twice-modified Lagrangian (16), we in place of (14.2) obtain

$$\begin{aligned} \partial_\mu F^{\mu\nu} &= \frac{\partial}{\partial A_\mu}\mathcal{L}_{\text{int}}(\psi, \psi^*, \partial\psi, \partial\psi^*, A) \\ &= \tfrac{1}{\mathcal{C}} J^\nu \end{aligned} \tag{17}$$

with

$$\begin{aligned} J^\nu &\equiv g\hbar c\big\{Q^\nu + g\tfrac{\hbar}{m}(\psi^*\psi)A^\nu\big\} \\ &= j^\nu + \tfrac{e^2}{mc}\psi^* A^\nu \psi \qquad (18.1) \\ j^\nu &\equiv eQ^\nu = J\Big|_{A\to 0} \qquad (18.2) \end{aligned}$$

The gauge-invariance of J^ν—required for the self-consistency of (17)—is not obvious (certainly not "manifest"), but is readily established.

The field equation (14.1) can be written

$$g^{\alpha\beta}\big(\tfrac{\hbar}{i}\partial_\alpha - \tfrac{e}{\mathcal{C}}A_\alpha\big)\big(\tfrac{\hbar}{i}\partial_\beta - \tfrac{e}{\mathcal{C}}A_\beta\big)\psi = (mc)^2\psi \tag{19}$$

and in this form can be considered to have resulted by ordinary Schrödinger quantization from a classical process of the form

$$\left.\begin{aligned} g^{\alpha\beta}p_\alpha p_\beta &= (mc)^2 \\ &\downarrow \\ g^{\alpha\beta}(p_\alpha - \tfrac{e}{\mathcal{C}}A_\alpha)(p_\beta - \tfrac{e}{\mathcal{C}}A_\beta) &= (mc)^2 \end{aligned}\right\} \tag{20}$$

Note also that

$$[j^\nu] = [J^\nu] = \frac{\text{electrical charge}}{\text{area}\cdot\text{time}} = \text{electrical current density}$$

and that we are now in position to write

$$\begin{aligned} \mathcal{L}_{\text{int}} &= \tfrac{1}{\mathcal{C}} j^\alpha A_\alpha + \underbrace{\tfrac{e^2}{2mc^2}(\psi^* A^\alpha \psi)}_{=\frac{1}{2}(J^\alpha - j^\alpha)} A_\alpha \\ &= \tfrac{1}{\mathcal{C}}\tfrac{1}{2}(J^\alpha + j^\alpha)A_\alpha \end{aligned} \tag{21}$$

Assuredly,

$$\text{local gauge invariance} \quad \Longrightarrow \quad \text{global gauge invariance}$$

and from the manifest invariance of the twice-modified Lagrangian (16) under the global instance

$$\begin{aligned} \psi &\longrightarrow \psi' = e^{+i\omega} \cdot \psi \\ A_\mu &\longrightarrow A'_\mu = A_\mu \end{aligned}$$

of (8)—which infinitesimally becomes

$$\begin{aligned} \psi &\longrightarrow \psi + \delta\psi \quad \text{with} \quad \delta\psi = +i\psi \cdot \delta\omega \\ \psi^* &\longrightarrow \psi^* + \delta\psi^* \quad \text{with} \quad \delta\psi^* = -i\psi \cdot \delta\omega \\ A_\mu &\longrightarrow A_\mu + \delta A_\mu \quad \text{with} \quad \delta A_\mu = 0 \end{aligned}$$

—we are, by Noether's theorem, led to the conservation of

$$\begin{aligned} -(e/\hbar)\left\{ \frac{\partial \mathcal{L}}{\partial \psi_{,\nu}}(+i\psi) + \frac{\partial \mathcal{L}}{\partial \psi^*_{,\nu}}(-i\psi^*) \right\} &= j^\nu - (e/\hbar)\left\{ \frac{\partial \mathcal{L}_{\text{int}}}{\partial \psi_{,\nu}}(i\psi) - \frac{\partial \mathcal{L}_{\text{int}}}{\partial \psi^*_{,\nu}}(i\psi^*) \right\} \\ &= j^\nu + \tfrac{e^2}{mc}\psi^* A^\nu \psi \\ &= J^\nu \end{aligned}$$

(i.e., to $\partial_\mu J^\mu = 0$) which at (17) was presented as an automatic consequence of the antisymmetry of $F^{\mu\nu}$. Prior to introduction of the gauge field A_μ we had $\partial_\mu j^\mu = 0$. We can in this light understand the adjustment

$$j^\nu(\psi, \psi^*, \partial\psi, \partial\psi^*) \quad \longrightarrow \quad J^\nu(\psi, \psi^*, \partial\psi, \partial\psi^*, A)$$

as a price paid in our effort

$$\mathcal{L}_0(\psi, \psi^*, \partial\psi, \partial\psi^*) \quad \longrightarrow \quad \mathcal{L}_2(\psi, \psi^*, \partial\psi, \partial\psi^*, A, \partial A)$$

to achieve local gauge invariance.

The effort to which I have just referred has yield up (amongst others) the equations

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad \text{and} \quad \partial_\mu F^{\mu\nu} = \tfrac{1}{c} J^\nu$$

It has, in short, delivered Maxwellian electrodynamics to us on a platter. If we were to "turn off" the ψ-field (or set the coupling constant $e = 0$) we would retain free-field electrodynamics as a kind of residue—a gift, for having shopped in the Gauge Store. The "compensating field" has been found in this instance to lead a busy physical life of its own, even when it has nothing to compensate. It was this development which first suggested that gauge field theory might, in fact, be good for something.

We have been supplied, moreover, with a detailed account of how the electromagnetic field $F^{\mu\nu}$ and the ψ-field are coupled—a "theory of field

interactions." It was, in fact, a quest for a general theory of field interactions which led Ronald Shaw—in 1953/54 a graduate student of Abdus Salam at Cambridge, working under the influence of Schwinger to the invention of gauge field theory, independently of (and almost simultaneously with) Yang & Mills.[8]

The electromagnetic aspects of the theory to which we have been led do, however, present one problematic (or at least surprising) aspect: the current term J^ν which "stimulates" the electromagnetic field was found at (18) to *itself depend upon* the field (through the 4-potential A_μ). As we move farther into our subject we will remain on the alert for developments which may serve to clarify that circumstance.

Gauge theory of a non-relativistic classical particle. Gauge field theory was born of general/special relativistic parents, and has spent its adult life married to quantum mechanics. It may be well, therefore, to be reminded that the central idea is so robust that it can flourish even when deprived of either or both of those controlling influences. To illustrate the point, I look to the classical Hamilton-Jacobi theory of a non-relativistic particle:

Assume the Hamiltonian to have the form $H(\boldsymbol{p},\boldsymbol{x}) = \frac{1}{2m}\boldsymbol{p}\cdot\boldsymbol{p} + U(\boldsymbol{x})$. The Hamilton-Jacobi equation then reads

$$\tfrac{1}{2m}\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S + U(\boldsymbol{x}) + S_t = 0 \tag{22.1}$$

and[9] when joined by its companion

$$R_t + \boldsymbol{\nabla}\cdot(\tfrac{1}{m}R\boldsymbol{\nabla}S) = 0 \tag{22.2}$$

can be consider to derive from the Lagrangian

$$\mathcal{L}_0(S,\partial S,R) = R\cdot\left\{\tfrac{1}{2m}\left[(\partial_x S)^2 + (\partial_y S)^2 + (\partial_z S)^2\right] + U + (\partial_t S)\right\} \tag{23}$$

This Lagrangian is manifestly invariant under the global gauge transformations

$$\left.\begin{array}{l} S \longrightarrow S' = S + \hbar\omega \\ R \longrightarrow R' = R \end{array}\right\} \tag{24}$$

which—compare (5)—are, in Pauli's terminology, "gauge transformations of the 2^{nd} kind," into which $\hbar$ has been introduced as a "soft constant of the action" in order to preserve the dimensionlessness of ω. To achieve the

$$\text{global gauge} \quad \Longrightarrow \quad \text{local gauge}$$

[8] Relevant sections of Shaw's thesis (1955) are reprinted in O'Raifeartaigh.[3]
[9] See again the discussion which culminated in (1–108).

symmetry enhancement we adjust the design of the system

$$\begin{aligned}
&\mathcal{L}_0(S,\partial S,R)\\
&\qquad\downarrow\\
&\mathcal{L}_0(S,\mathcal{D}S,R) = R\cdot\Big\{\tfrac{1}{2m}\big[(\partial_x S+g\hbar C_x)^2+(\partial_y S+g\hbar C_y)^2+(\partial_z S+g\hbar C_z)^2\big]\\
&\qquad\qquad\qquad\qquad\qquad\qquad +U+(\partial_t S+g\hbar C_t)\Big\}\\
&\qquad\qquad\equiv \mathcal{L}_1(S,\partial S,R,C_x,C_y,C_z,C_t)
\end{aligned}\tag{25}$$

and—so as to achieve

$$\partial_x S' + g\hbar C'_x = \partial_x S + g\hbar C_x \ , \ \text{etc.}$$

—assign to the notion of a "local gauge transformation" an enlarged meaning

$$\left.\begin{aligned}
S &\longrightarrow S' = S + g\hbar\Omega(\boldsymbol{x},t)\\
R &\longrightarrow R' = R\\
C_x &\longrightarrow C'_x = C_x - \partial_x\Omega(\boldsymbol{x},t)\\
C_y &\longrightarrow C'_y = C_y - \partial_y\Omega(\boldsymbol{x},t)\\
C_z &\longrightarrow C'_z = C_z - \partial_y\Omega(\boldsymbol{x},t)\\
C_t &\longrightarrow C'_t = C_t - \partial_t\,\Omega(\boldsymbol{x},t)
\end{aligned}\right\}\tag{26}$$

which in the global case $g\Omega(\boldsymbol{x},t)=\omega$ gives back essentially (24).

I turn now to remarks intended to help us mix some gauge-invariant ∂C-dependence into (25), and thus to lauch the gauge fields C into dynamical motion. We proceed from the observation that the expressions[10]

$$\begin{array}{lll}
\tfrac{1}{c}(\partial_t C_x - \partial_x C_t) & \tfrac{1}{c}(\partial_t C_y - \partial_y C_t) & \tfrac{1}{c}(\partial_t C_z - \partial_z C_t)\\
 & (\partial_x C_y - \partial_y C_x) & (\partial_x C_z - \partial_z C_x)\\
 & & (\partial_y C_z - \partial_z C_y)
\end{array}$$

are individually gauge-invariant (because the cross partials of Ω are equal). To keep our theory from coming *rotationally* unstuck, we must require that

$$\begin{pmatrix} C_x\\ C_y\\ C_z\end{pmatrix} \text{ transforms like } \boldsymbol{\nabla}S\text{; i.e., as a vector}$$

Let us agree to write

$$\begin{pmatrix} C_x\\ C_y\\ C_z\end{pmatrix} \equiv -\boldsymbol{A} \quad\text{and}\quad C_t \equiv c\phi$$

[10] Some "soft c" factors have been introduced so as to render all entries co-dimensional.

where a sign has been introduced to establish contact with pre-established convention. The preceding tableau then becomes

$$\begin{array}{lll} (-\frac{1}{c}\partial_t \boldsymbol{A} - \boldsymbol{\nabla}\phi)_x & (-\frac{1}{c}\partial_t \boldsymbol{A} - \boldsymbol{\nabla}\phi)_y & (-\frac{1}{c}\partial_t \boldsymbol{A} - \boldsymbol{\nabla}\phi)_z \\ & +(\boldsymbol{\nabla}\times\boldsymbol{A})_z & -(\boldsymbol{\nabla}\times\boldsymbol{A})_y \\ & & +(\boldsymbol{\nabla}\times\boldsymbol{A})_x \end{array}$$

which we will agree to abbreviate

$$\begin{array}{lll} (\boldsymbol{E})_x & (\boldsymbol{E})_y & (\boldsymbol{E})_z \\ & -(\boldsymbol{B})_z & +(\boldsymbol{B})_y \\ & & -(\boldsymbol{B})_x \end{array}$$

The expressions $\boldsymbol{E}\cdot\boldsymbol{E}$, $\boldsymbol{E}\cdot\boldsymbol{B}$ and $\boldsymbol{B}\cdot\boldsymbol{B}$ are

- quadratic in ∂C
- gauge-invariant, and
- rotationally invariant

and candidates, therefore, for independent inclusion into the design of a modified Lagrangian. Our most recent Lagrangian (25) can, in present notation, be written[11]

$$\mathcal{L}_1 = R\cdot\Big\{\tfrac{1}{2m}\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\cdot\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big) + U + (\partial_t S + e\phi)\Big\} \tag{27}$$

and we are led by the preceding remarks to consider Lagrangians of the modified form

$$\begin{aligned} \mathcal{L}_2(S,\partial S,R,\boldsymbol{A},\phi,\partial\boldsymbol{A},\partial\phi) &= \mathcal{L}_1 + \tfrac{1}{2}p\boldsymbol{E}\cdot\boldsymbol{E} + q\boldsymbol{E}\cdot\boldsymbol{B} + \tfrac{1}{2}r\boldsymbol{B}\cdot\boldsymbol{B} \\ &= \mathcal{L}_1 + \tfrac{1}{2}p(\tfrac{1}{c}\partial_t\boldsymbol{A} + \boldsymbol{\nabla}\phi)\cdot(\tfrac{1}{c}\partial_t\boldsymbol{A} + \boldsymbol{\nabla}\phi) \\ &\quad - \ q(\tfrac{1}{c}\partial_t\boldsymbol{A} + \boldsymbol{\nabla}\phi)\cdot(\boldsymbol{\nabla}\times\boldsymbol{A}) \\ &\quad + \tfrac{1}{2}r(\boldsymbol{\nabla}\times\boldsymbol{A})\cdot(\boldsymbol{\nabla}\times\boldsymbol{A}) \end{aligned} \tag{28}$$

Look to the associated field equations and conservation laws. From

$$\Big\{\partial_t\frac{\partial}{\partial R_t} + \boldsymbol{\nabla}\cdot\frac{\partial}{\partial\boldsymbol{\nabla}R} - \frac{\partial}{\partial R}\Big\}\mathcal{L}_2 = 0$$

we obtain[12]

$$\tfrac{1}{2m}\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\cdot\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big) + U + (\partial_t S + e\phi) = 0 \tag{28.1}$$

[11] One needs to notice that the compensating fields C have—by contrivance—the same physical dimension as the gauge fields formerly notated A_μ (to which they are really identical), and that consequently it still makes sense to write $g\hbar = e/c$.

[12] Compare this with the $H = \frac{1}{2m}(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A})\cdot(\boldsymbol{p} - \frac{e}{c}\boldsymbol{A}) + e\phi$ which appears, for example, as (8–27) in Goldstein's *Classical Mechanics* (2nd edition 1980). There it arises from $L = \frac{1}{2}m\dot{\boldsymbol{x}}\cdot\dot{\boldsymbol{x}} + \frac{e}{c}\dot{\boldsymbol{x}}\cdot\boldsymbol{A} - e\phi$, which hinges on the observation that the Lorentz force law $\boldsymbol{F} = e(\boldsymbol{E} + \frac{1}{c}\dot{\boldsymbol{x}}\times\boldsymbol{B}) = e\big\{-\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial}{\partial t}\boldsymbol{A} + \frac{1}{c}\dot{\boldsymbol{x}}\times\boldsymbol{\nabla}\times\boldsymbol{A}\big\}$ can be obtained by Lagrange differentiation of $U = e\big(\phi - \frac{1}{c}\dot{\boldsymbol{x}}\cdot\boldsymbol{A}\big)$. See Goldstein's §1–5 for details.

while

$$\Big\{\partial_t \frac{\partial}{\partial S_t} + \boldsymbol{\nabla}\cdot \frac{\partial}{\partial \boldsymbol{\nabla} S} - \frac{\partial}{\partial S}\Big\}\mathcal{L}_2 = 0$$

gives

$$\partial_t R + \boldsymbol{\nabla}\cdot \tfrac{1}{m} R\big(\boldsymbol{\nabla} S - \tfrac{e}{c}\boldsymbol{A}\big) = 0 \tag{28.2}$$

Interestingly, this last field equation displays the design of a continuity equation, and is in fact *precisely conservation law which by Noether's theorem arises as an expression of the global gauge invariance* of the locally gauge-invariant Lagrangian (27). A simple dimensional argument gives $[R] = 1/(\text{length})^3$, so if we introduce the notations

$$\left.\begin{array}{lcl} \rho \equiv eR & : & \text{charge density} \\ \boldsymbol{J} \equiv \frac{e}{m} R\big(\boldsymbol{\nabla} S - \frac{e}{c}\boldsymbol{A}\big) & : & \text{current density} \end{array}\right\} \tag{29}$$

then (28.2) can be read as a statement

$$\partial_t \rho + \boldsymbol{\nabla}\cdot\boldsymbol{J} = 0 \tag{30}$$

of charge conservation. Looking finally to the dynamics of the gauge fields: from

$$\Big\{\partial_t \frac{\partial}{\partial \phi_t} + \boldsymbol{\nabla}\cdot \frac{\partial}{\partial \boldsymbol{\nabla}\phi} - \frac{\partial}{\partial \phi}\Big\}\mathcal{L}_2 = 0$$

we obtain $\boldsymbol{\nabla}\cdot\big\{p(\frac{1}{c}\partial_t \boldsymbol{A} + \boldsymbol{\nabla}\phi) - q\boldsymbol{\nabla}\times\boldsymbol{A}\big\} - eR = 0$ which (recall that div curl always vanishes) can be expressed $-p\boldsymbol{\nabla}\cdot\boldsymbol{E} = \rho$, while

$$\Big\{\partial_t \frac{\partial}{\partial A_{x,t}} + \partial_x \frac{\partial}{\partial A_{x,x}} + \partial_y \frac{\partial}{\partial A_{x,y}} + \partial_z \frac{\partial}{\partial A_{x,z}} - \frac{\partial}{\partial A_x}\Big\}\mathcal{L}_2 = 0$$

is found after simplifications to yield the x-component of

$$p\tfrac{1}{c}\partial_t \boldsymbol{E} + q\big(\tfrac{1}{c}\partial_t \boldsymbol{B} + \boldsymbol{\nabla}\times\boldsymbol{E}\big) + r\boldsymbol{\nabla}\times\boldsymbol{B} = \tfrac{1}{c}\boldsymbol{J}$$

But from the standing definitions

$$\boldsymbol{E} \equiv -\tfrac{1}{c}\partial_t \boldsymbol{A} - \boldsymbol{\nabla}\phi \quad \text{and} \quad \boldsymbol{B} \equiv \boldsymbol{\nabla}\times\boldsymbol{A} \tag{31}$$

it follows automatically that

$$\left.\begin{array}{r} \boldsymbol{\nabla}\cdot\boldsymbol{B} = 0 \\ \frac{1}{c}\partial_t \boldsymbol{B} + \boldsymbol{\nabla}\times\boldsymbol{E} = \boldsymbol{0} \end{array}\right\} \tag{32.1}$$

and from $\mathcal{L}_2$ we have obtained this additional information:

$$\left.\begin{array}{r} -p\boldsymbol{\nabla}\cdot\boldsymbol{E} = \rho \\ p\frac{1}{c}\partial_t \boldsymbol{E} + r\boldsymbol{\nabla}\times\boldsymbol{B} = \frac{1}{c}\boldsymbol{J} \end{array}\right\} \tag{32.2}$$

Note that (32.2) renders $\partial_t \rho + \boldsymbol{\nabla}\cdot\boldsymbol{J} = 0$ automatic in all cases (as it must, since the continuity equation derives from built-in global gauge invariance), and that equations (32.2)

become precisely the sourcey Maxwell equations in the case $r = -p = 1$

But what heretofore neglected physical principle serves to enforce such conditions? An answer emerges from study of the energy/momentum/angular momentum properties of the gauged Hamilton-Jacobi field system.

By way of preparation, we look first to those properties as they refer to the ungauged system $\mathcal{L}_0$ which provided our point of departure. A little exploratory tinkering motivates these definitions:

$$\begin{aligned}
\mathcal{E} &\equiv -\Big[\Big\{R_t\frac{\partial}{\partial R_t} + S_t\frac{\partial}{\partial S_t}\Big\}\mathcal{L}_0 - \mathcal{L}_0\Big] && (33.10)\\
&= R\cdot\Big\{\tfrac{1}{2m}\big[(\partial_x S)^2 + (\partial_y S)^2 + (\partial_z S)^2\big] + U\Big\}\\
&= R\cdot H(\boldsymbol{\nabla}S, \boldsymbol{x}) && (33.11)
\end{aligned}$$

$$\begin{aligned}
\mathcal{F}^x &\equiv -\Big[\Big\{R_t\frac{\partial}{\partial R_x} + S_t\frac{\partial}{\partial S_x}\Big\}\mathcal{L}_0\Big] && (33.20)\\
&= -\big[R\cdot\tfrac{1}{m}(\partial_x S)\big]S_t, \quad \text{with } \mathcal{F}^y \text{ and } \mathcal{F}^z \text{ described similarly}\\
&= +\big[R\cdot\tfrac{1}{m}(\partial_x S)\big]\cdot H(\boldsymbol{\nabla}S, \boldsymbol{x}) \quad \text{by the Hamilton-Jacobi equation}\\
&\downarrow\\
\boldsymbol{\mathcal{F}} &= \big[\tfrac{1}{m}R\,\boldsymbol{\nabla}S\big]\cdot H(\boldsymbol{\nabla}S, \boldsymbol{x}) && (33.21)
\end{aligned}$$

$$\begin{aligned}
\mathcal{P}_x &\equiv +\Big[\Big\{R_x\frac{\partial}{\partial R_t} + S_x\frac{\partial}{\partial S_t}\Big\}\mathcal{L}_0\Big] && (33.30)\\
&= R\cdot(\partial_x S), \quad \text{with } \mathcal{P}_y \text{ and } \mathcal{P}_z \text{ described similarly}\\
&\downarrow\\
\boldsymbol{\mathcal{P}} &= R\,\boldsymbol{\nabla}S && (33.31)
\end{aligned}$$

$$\begin{aligned}
\mathcal{T}^u{}_v &\equiv +\Big[\Big\{R_v\frac{\partial}{\partial R_u} + S_v\frac{\partial}{\partial S_u}\Big\}\mathcal{L}_0 - \delta^u{}_v\mathcal{L}_0\Big] \;:\; \big\{u,v\big\} \in \big\{x,y,z\big\} && (33.40)\\
&= R\cdot\big[\tfrac{1}{m}S_u S_v - H\delta^u{}_v\big] && (33.41)
\end{aligned}$$

It is not difficult to establish that

$$\begin{aligned}
\tfrac{\partial}{\partial t}\mathcal{E} + \boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{F}} &= +R\,\tfrac{\partial}{\partial t}U && (34.1)\\
&= 0 \quad \text{if } U \text{ is } t\text{-independent}
\end{aligned}$$

$$\begin{aligned}
\tfrac{\partial}{\partial t}\mathcal{P}_x + \tfrac{\partial}{\partial x}\mathcal{T}^x{}_x + \tfrac{\partial}{\partial y}\mathcal{T}^y{}_x + \tfrac{\partial}{\partial z}\mathcal{T}^z{}_x &= -R\,\tfrac{\partial}{\partial x}U \;:\; \text{ditto with } {}_x \to {}_{y,z} && (34.2)\\
&= 0 \quad \text{if } U \text{ is } x\text{-independent}
\end{aligned}$$

and from the manifest *symmetry* of ${\mathcal{T}^u}_v$ it follows that angular momentum is locally conserved at points where the "torque density" $\boldsymbol{x} \times (-\boldsymbol{\nabla} U)$ vanishes. These are physically satisfying results, but my main point has been to identify the contrasting *signs* which enter most sensibly into the preceding definitions.

Preserving those sign conventions, we look now to the energy/momentum desnities/fluxes which arise from

$$\begin{aligned} \mathcal{L}_{\text{free gauge field}} &= \tfrac{1}{2}p(\tfrac{1}{c}\partial_t \boldsymbol{A} + \boldsymbol{\nabla}\phi)\cdot(\tfrac{1}{c}\partial_t \boldsymbol{A} + \boldsymbol{\nabla}\phi) + \tfrac{1}{2}r\,(\boldsymbol{\nabla}\times\boldsymbol{A})\cdot(\boldsymbol{\nabla}\times\boldsymbol{A}) \\ &= \tfrac{1}{2}p\,\boldsymbol{E}\cdot\boldsymbol{E} + \tfrac{1}{2}r\boldsymbol{B}\cdot\boldsymbol{B} \end{aligned} \tag{35}$$

where the originally conjectured q-term, since it made no contribution to the field equations (32.2), has been dropped, and where it is for present computational purposes most efficient to write

$$\mathcal{L}_{\text{free gauge field}} = -p\boldsymbol{E}\cdot\begin{pmatrix} \frac{1}{c}A_{x,t} + \phi_{,x} \\ \frac{1}{c}A_{y,t} + \phi_{,y} \\ \frac{1}{c}A_{z,t} + \phi_{,z} \end{pmatrix} + r\boldsymbol{B}\cdot\begin{pmatrix} A_{z,y} - A_{y,z} \\ A_{x,z} - A_{z,x} \\ A_{y,x} - A_{x,y} \end{pmatrix}$$

Looking first to the *energy density* of the free gauge field system, we find

$$\begin{aligned} \tilde{\mathcal{E}} &\equiv -\Big[\Big\{\phi_{,t}\frac{\partial}{\partial\phi_{,t}} + A_{x,t}\frac{\partial}{\partial A_{x,t}} + A_{y,t}\frac{\partial}{\partial A_{y,t}} + A_{z,t}\frac{\partial}{\partial A_{z,t}}\Big\} - 1\Big]\mathcal{L}_{\text{gauge}} \\ &= -\big[-p\tfrac{1}{c}\boldsymbol{A}_t\cdot\boldsymbol{E} - \tfrac{1}{2}p\,\boldsymbol{E}\cdot\boldsymbol{E} - \tfrac{1}{2}r\boldsymbol{B}\cdot\boldsymbol{B}\big] \\ &= \underbrace{-\tfrac{1}{2}p\,\boldsymbol{E}\cdot\boldsymbol{E} + \tfrac{1}{2}r\boldsymbol{B}\cdot\boldsymbol{B}}_{\mathcal{E}_{\text{gauge}}} \underbrace{-\,p\,\boldsymbol{E}\cdot\boldsymbol{\nabla}\phi}_{\text{gauge-dependent term, soon discarded}} \quad \text{by} \quad -\tfrac{1}{c}\partial_t\boldsymbol{A} = \boldsymbol{E} + \boldsymbol{\nabla}\phi \end{aligned} \tag{36.1}$$

and notice that $\mathcal{E} \geqslant 0$ requires $p < 0$ and $r > 0$. Without loss of generality (since the numerical part of p can be absorbed into the definition of e) we

$$\text{Set } p = -1$$

whereupon (32.2) become

$$\left.\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{E} &= \rho \\ r\boldsymbol{\nabla}\times\boldsymbol{B} &= \tfrac{1}{c}\big\{\boldsymbol{J} + \tfrac{1}{c}\partial_t\boldsymbol{E}\big\} \end{aligned}\right\} \tag{32.2$'$}$$

For the *components of energy flux* we have

$$\begin{aligned} \tilde{\mathcal{F}}^x &\equiv -\Big[\phi_{,t}\frac{\partial}{\partial\phi_{,x}} + A_{x,t}\frac{\partial}{\partial A_{x,x}} + A_{y,t}\frac{\partial}{\partial A_{y,x}} + A_{z,t}\frac{\partial}{\partial A_{z,x}}\Big]\mathcal{L}_{\text{gauge}} \\ &= -\big[\phi_t E_x + rA_{y,t}B_z - rA_{z,t}B_y\big], \quad \text{with } \tilde{\mathcal{F}}^y \text{ and } \tilde{\mathcal{F}}^z \text{ described similarly} \\ &\downarrow \\ \tilde{\boldsymbol{\mathcal{F}}} &= -\big[r(\partial_t\boldsymbol{A})\times\boldsymbol{B} + \boldsymbol{E}\,\partial_t\phi\big] \\ &= \underbrace{rc(\boldsymbol{E}\times\boldsymbol{B})}_{\boldsymbol{\mathcal{F}}_{\text{gauge}}} + \underbrace{\big\{rc\boldsymbol{\nabla}\phi\times\boldsymbol{B} - \boldsymbol{E}\,\partial_t\phi\big\}}_{\text{gauge-dependent term, soon discarded}} \end{aligned} \tag{36.2}$$

The "gauge-dependent terms" which enter additively into equations (36) cannot participate in the physical output of the theory, but from results already in hand it follows readily that

$$\begin{aligned}\tfrac{\partial}{\partial t}(\boldsymbol{E}\cdot\boldsymbol{\nabla}\phi)+\boldsymbol{\nabla}\cdot\big\{rc\boldsymbol{\nabla}\phi\times\boldsymbol{B}-\boldsymbol{E}\,\partial_t\phi\big\}&=-\big\{\boldsymbol{J}\cdot\boldsymbol{\nabla}+\rho\partial_t\big\}\phi \\ &=0\quad\text{in the absence of sources}\end{aligned}\tag{37}$$

so those terms can be discarded on grounds that they make no contribution to the *total* energy resident in the gauge field system. Notice that r remains still indeterminate.

Looking next to the *components of momentum density* in the free gauge field system, we have

$$\begin{aligned}\tilde{\mathcal{P}}_x&\equiv\Big[\phi_{,x}\frac{\partial}{\partial\phi_{,t}}+A_{x,x}\frac{\partial}{\partial A_{x,t}}+A_{y,x}\frac{\partial}{\partial A_{y,t}}+A_{z,x}\frac{\partial}{\partial A_{z,t}}\Big]\mathcal{L}_{\text{gauge}}\\&=\tfrac{1}{c}\big[\boldsymbol{E}\cdot(\partial_x\boldsymbol{A})\big],\quad\text{with }\tilde{\mathcal{P}}_y\text{ and }\tilde{\mathcal{P}}_z\text{ described similarly}\end{aligned}$$

which (by a seldom-encountered but easily established identity) yields

$$\begin{aligned}\tilde{\boldsymbol{\mathcal{P}}}=\tfrac{1}{c}\begin{pmatrix}\boldsymbol{E}\cdot(\partial_x\boldsymbol{A})\\ \boldsymbol{E}\cdot(\partial_y\boldsymbol{A})\\ \boldsymbol{E}\cdot(\partial_z\boldsymbol{A})\end{pmatrix}&=\tfrac{1}{c}\boldsymbol{E}\times(\boldsymbol{\nabla}\times\boldsymbol{A})+\tfrac{1}{c}(\boldsymbol{E}\cdot\boldsymbol{\nabla})\boldsymbol{A}\\&=\underbrace{\tfrac{1}{c}\boldsymbol{E}\times\boldsymbol{B}}_{\boldsymbol{\mathcal{P}}_{\text{gauge}}}+\text{gauge-dependent term}\end{aligned}\tag{38.1}$$

Looking finally to the *components of the* 3×3 *stress tensor* (i.e., of momentum flux), we use

$$\tilde{\mathcal{T}}^{u}{}_{v}=\Big[\phi_{,v}\frac{\partial}{\partial\phi_{,u}}+A_{x,v}\frac{\partial}{\partial A_{x,u}}+A_{y,v}\frac{\partial}{\partial A_{y,u}}+A_{z,v}\frac{\partial}{\partial A_{z,u}}-\delta^{u}{}_{v}\Big]\mathcal{L}_{\text{gauge}}$$

and writing

$$\|\tilde{\mathcal{T}}^{u}{}_{v}\|=\left(\begin{pmatrix}\tilde{\mathcal{T}}^{x}{}_{x}\\ \tilde{\mathcal{T}}^{y}{}_{x}\\ \tilde{\mathcal{T}}^{z}{}_{x}\end{pmatrix}\quad\begin{pmatrix}\tilde{\mathcal{T}}^{x}{}_{y}\\ \tilde{\mathcal{T}}^{y}{}_{y}\\ \tilde{\mathcal{T}}^{z}{}_{y}\end{pmatrix}\quad\begin{pmatrix}\tilde{\mathcal{T}}^{x}{}_{z}\\ \tilde{\mathcal{T}}^{y}{}_{z}\\ \tilde{\mathcal{T}}^{z}{}_{z}\end{pmatrix}\right)$$

compute

$$\begin{pmatrix}\tilde{\mathcal{T}}^{x}{}_{x}\\ \tilde{\mathcal{T}}^{y}{}_{x}\\ \tilde{\mathcal{T}}^{z}{}_{x}\end{pmatrix}=\boldsymbol{E}\,(\partial_x\phi)+r(\partial_x\boldsymbol{A})\times\boldsymbol{B}+\big\{\tfrac{1}{2}\,\boldsymbol{E}\cdot\boldsymbol{E}-\tfrac{1}{2}r\boldsymbol{B}\cdot\boldsymbol{B}\big\}\begin{pmatrix}1\\0\\0\end{pmatrix},\text{etc.}$$

which (by a seldom-encountered and not-so-easily established population of identities—see below) yields

$$= -\boldsymbol{E}E_x - r\boldsymbol{B}B_x + \left\{\tfrac{1}{2}\,\boldsymbol{E}\cdot\boldsymbol{E} + \tfrac{1}{2}r\boldsymbol{B}\cdot\boldsymbol{B}\right\}\begin{pmatrix}1\\0\\0\end{pmatrix} - \tfrac{1}{c}\boldsymbol{E}(\partial_t A_x) - r(\boldsymbol{B}\times\boldsymbol{\nabla})A_x$$

$$= \begin{pmatrix}\mathcal{T}^x{}_x\\ \mathcal{T}^y{}_x\\ \mathcal{T}^z{}_x\end{pmatrix}_{\text{gauge}} + \underbrace{\left\{-\tfrac{1}{c}\boldsymbol{E}(\partial_t A_x) - r(\boldsymbol{B}\times\boldsymbol{\nabla})A_x\right\}}_{\text{gauge-dependent term}} \tag{38.2}$$

with

$$\begin{pmatrix}\mathcal{T}^x{}_x\\ \mathcal{T}^y{}_x\\ \mathcal{T}^z{}_x\end{pmatrix}_{\text{gauge}} \equiv \begin{pmatrix}-E_xE_x - rB_xB_x + \left\{\tfrac{1}{2}\,\boldsymbol{E}\cdot\boldsymbol{E} + \tfrac{1}{2}r\boldsymbol{B}\cdot\boldsymbol{B}\right\}\\ -E_yE_x - rB_yB_x\\ -E_zE_x - rB_zB_x\end{pmatrix}$$

But—looking now more closely to the gauge-dependent terms—we find

$$\begin{aligned}
&\tfrac{\partial}{\partial t}\left\{\tfrac{1}{c}(\boldsymbol{E}\cdot\boldsymbol{\nabla})A_x\right\} + \boldsymbol{\nabla}\cdot\left\{-\tfrac{1}{c}\boldsymbol{E}(\partial_t A_x) - r(\boldsymbol{B}\times\boldsymbol{\nabla})A_x\right\}\\
&\quad= \left\{-\tfrac{1}{c}\boldsymbol{j} + r\boldsymbol{\nabla}\times\boldsymbol{B}\right\}\cdot\nabla A_x + \tfrac{1}{c}\boldsymbol{E}\cdot\boldsymbol{\nabla}(\partial_t A_x)\\
&\qquad\qquad - \tfrac{1}{c}\boldsymbol{E}\cdot\boldsymbol{\nabla}(\partial_t A_x) - \tfrac{1}{c}\rho(\partial_t A_x) - r\boldsymbol{\nabla}\cdot(\boldsymbol{B}\times\boldsymbol{\nabla})A_x\\
&\quad= -\tfrac{1}{c}\left\{\boldsymbol{J}\cdot\boldsymbol{\nabla} + \rho\partial_t\right\}A_x + r\underbrace{\left\{\boldsymbol{\nabla}\times\boldsymbol{B}\cdot\nabla - \boldsymbol{\nabla}\cdot(\boldsymbol{B}\times\boldsymbol{\nabla})\right\}}_{0,\ \text{by quick demonstration}}A_x\\
&\quad= 0 \quad \text{in the absence of sources}
\end{aligned} \tag{39}$$

This result supplies the familiar grounds on which we will *abandon* the gauge-sensitive terms. Postponing discussion of the results now in hand...

I digress now to establish the identity

$$\begin{pmatrix}\left[(\partial_x\boldsymbol{A})\times\boldsymbol{B}\right]_x\\ \left[(\partial_x\boldsymbol{A})\times\boldsymbol{B}\right]_y\\ \left[(\partial_x\boldsymbol{A})\times\boldsymbol{B}\right]_z\end{pmatrix} = \begin{pmatrix}-B_xB_x + \boldsymbol{B}\cdot\boldsymbol{B}\\ -B_yB_x\\ -B_zB_x\end{pmatrix} - \begin{pmatrix}(\boldsymbol{B}\times\boldsymbol{\nabla})_xA_x\\ (\boldsymbol{B}\times\boldsymbol{\nabla})_yA_x\\ (\boldsymbol{B}\times\boldsymbol{\nabla})_zA_x\end{pmatrix} \tag{40}$$

used in the argument which led to (38.2). By way of preparation, we note that

$$\boldsymbol{B}\times\boldsymbol{B} = \boldsymbol{B}\times(\boldsymbol{\nabla}\times\boldsymbol{A}) = \boldsymbol{0} \quad\Rightarrow\quad \begin{cases}B_yA_{y,x} - B_yA_{y,x} = B_zA_{x,z} - B_zA_{z,x}\\ B_zA_{z,y} - B_zA_{y,z} = B_xA_{y,x} - B_xA_{x,y}\\ B_xA_{x,z} - B_xA_{z,x} = B_yA_{z,y} - B_yA_{y,z}\end{cases}$$

and that

$$\boldsymbol{B}\cdot\boldsymbol{B} = B_x(A_{z,y} - A_{z,y}) + B_y(A_{x,z} - A_{z,x}) + B_z(A_{y,x} - A_{x,y})$$

Drawing without specific comment upon those facts, we have

$$\begin{aligned}
\big[(\partial_x \boldsymbol{A})\times\boldsymbol{B}\big]_x &= B_z A_{y,x} - B_y A_{z,x} \\
&= -B_x B_x + \boldsymbol{B}\cdot\boldsymbol{B} + \big\{B_z A_{y,x} - B_y A_{z,x} - B_y B_y - B_z B_z\big\} \\
\big\{\text{etc.}\big\} &= B_z A_{y,x} - B_y A_{z,x} - B_y A_{x,z} + B_y A_{z,x} - B_z A_{y,x} + B_z A_{x,y} \\
&= -B_y A_{x,z} + B_z A_{x,y} \\
&= -(\boldsymbol{B}\times\boldsymbol{\nabla})_x A_x \\
\big[(\partial_x \boldsymbol{A})\times\boldsymbol{B}\big]_y &= B_x A_{z,x} - B_z A_{x,x} \\
&= -B_y B_x + \big\{B_x A_{z,x} - B_z A_{x,x} + B_y(A_{z,y} - A_{y,z})\big\} \\
\big\{\text{etc.}\big\} &= B_x A_{z,x} - B_z A_{x,x} + B_x(A_{x,z} - A_{z,x}) \\
&= -B_z A_{x,x} + B_x A_{x,z} \\
&= -(\boldsymbol{B}\times\boldsymbol{\nabla})_y A_x \\
\big[(\partial_x \boldsymbol{A})\times\boldsymbol{B}\big]_z &= B_y A_{x,x} - B_x A_{y,x} \\
&= -B_z B_x + \big\{B_y A_{x,x} - B_x A_{y,x} + B_z(A_{z,y} - A_{y,z})\big\} \\
\big\{\text{etc.}\big\} &= B_y A_{x,x} - B_x A_{y,x} + B_x(A_{y,x} - A_{x,y}) \\
&= -B_x A_{x,y} + B_y A_{x,x} \\
&= -(\boldsymbol{B}\times\boldsymbol{\nabla})_z A_x
\end{aligned}$$

which serve to establish the identity in question. Companion identities are obtained by cyclic permutation on $\big\{x, y, z\big\}$.

The results recently acquired are summarized in the following display:

$$\begin{pmatrix} \mathcal{E} & \mathcal{P}_x & \mathcal{P}_y & \mathcal{P}_z \\ \mathcal{F}^x & \mathcal{T}^x{}_x & \mathcal{T}^x{}_y & \mathcal{T}^x{}_z \\ \mathcal{F}^y & \mathcal{T}^y{}_x & \mathcal{T}^y{}_y & \mathcal{T}^y{}_z \\ \mathcal{F}^z & \mathcal{T}^z{}_x & \mathcal{T}^z{}_y & \mathcal{T}^z{}_z \end{pmatrix} = \begin{pmatrix} \mathcal{E} & \frac{1}{c}(\boldsymbol{E}\times\boldsymbol{B})^{\intercal} \\ rc\boldsymbol{E}\times\boldsymbol{B} & \mathbb{T} \end{pmatrix} \tag{41}$$

where $\mathcal{E} \equiv \frac{1}{2}(E^2 + rB^2)$ and

$$\mathbb{T} \equiv \begin{pmatrix} \mathcal{E} - E_x E_x - rB_x B_x & -E_x E_y - rB_x B_y & -E_x E_z - rB_x B_z \\ -E_y E_x - rB_y B_x & \mathcal{E} - E_y E_y - rB_y B_y & -E_y E_z - rB_y B_z \\ -E_z E_x - rB_z B_x & -E_z E_y - rB_z B_y & \mathcal{E} - E_z E_z - rB_z B_z \end{pmatrix}$$

These results (except, perhaps, for the intrusion of the r-factors, concerning which I will have more to say in a moment) are of precisely the design supplied by Maxwellian electrodynamics.[13] The symmetry $\mathbb{T}^{\intercal} = \mathbb{T}$ of the "stress tensor" is manifest (no Belinfante symmetrization was required after abandonment of the gauge-sensitive terms), and has been shown to assure *angular momentum conservation* in the free gauge field system.

[13] See D. Griffiths' *Introduction to Electrodynamics* (1981) §7.5; CLASSICAL ELECTRODYNAMICS (1980), p. 300.

We have been led from the non-relativistic classical mechanics written into the non-relativistic Hamilton-Jacobi Lagrangian (23) to the two-parameter class of gauge field theories implicit in the locally gauge-invariant Lagrangian (28) (from which we may consider the physically inconsequential q-term to have been dropped). A physical consideration ($\mathcal{E} \geqslant 0$) has led us to require

$$p < 0 < r$$

but has placed no restrictions upon the *numerical values* of p and r; we found it convenient at one point to set $p = -1$, but were certainly under no obligation to do so. Suppose we set $p = -s$ (with $s > 0$) and rewrite (32) as follows:

$$\left.\begin{aligned} \boldsymbol{\nabla}\cdot[\boldsymbol{B}/c] &= 0 \\ \boldsymbol{\nabla}\times\boldsymbol{E} &= -\partial_t[\boldsymbol{B}/c] \\ \boldsymbol{\nabla}\cdot\boldsymbol{E} &= \tfrac{1}{\epsilon}[\epsilon\rho/s] \\ \boldsymbol{\nabla}\times[\boldsymbol{B}/c] &= \tfrac{s}{rc^2\epsilon}[\epsilon\boldsymbol{J}/s] + \tfrac{s}{rc^2}\partial_t\boldsymbol{E} \\ &= \mu[\epsilon\boldsymbol{J}/s] + \mu\epsilon\partial_t\boldsymbol{E} \quad \text{with} \quad c^2\mu\epsilon \equiv s/r = -p/r \end{aligned}\right\} \tag{42}$$

Notational adjustments

$$[\boldsymbol{B}/c] \longrightarrow \boldsymbol{B}, \quad [\epsilon\rho/s] \longrightarrow \rho, \quad [\epsilon\boldsymbol{J}/s] \longrightarrow \boldsymbol{J}$$

lead then to "Maxwell equations" identical to those presented by Griffiths[13] at the beginning of his §7.3.3. In short: gauge theory has led us to a *population* of field theories, any one of which we are prepared by our experience to call "Maxwellian electrodynamics in an isotropic homogeneous medium." Each of those theories has a "relativistic look about it," but only one is relativistic in the Einsteinian sense—namely the one which results when (in effect) one sets

$$-p = r = 1 \quad \text{and} \quad c = \text{the observed constant of Nature}$$

In that case (35) reads

$$\begin{aligned} \mathcal{L}_{\text{free gauge field}} &= -\tfrac{1}{2}\,(\tfrac{1}{c}\partial_t\boldsymbol{A} + \boldsymbol{\nabla}\phi)\cdot(\tfrac{1}{c}\partial_t\boldsymbol{A} + \boldsymbol{\nabla}\phi) + \tfrac{1}{2}\,(\boldsymbol{\nabla}\times\boldsymbol{A})\cdot(\boldsymbol{\nabla}\times\boldsymbol{A}) \\ &= -\tfrac{1}{2}\,(\boldsymbol{E}\cdot\boldsymbol{E} - \boldsymbol{B}\cdot\boldsymbol{B}) \end{aligned} \tag{43}$$

which is familiar[14] as a Lorentz invariant

$$= -\tfrac{1}{4}F^{\alpha\beta}F_{\beta\alpha}$$

associated with the electromagnetic field *in vacuuo*. The take-home lesson: gauge theory is "relativistically predisposed," but does not *force* relativity upon us.

[14] See p. 256 in the notes just cited.

The discussion subsequent to (35) was off-puttingly dense. I might have made it less so by using sprinkled "it can be shown"s to surpress details (which I included because they frequently take surprising turns), but it seems well to recognize *why* the discussion was so cluttered: it was, in substantial part, because we worked non-relativistically—deprived of the organizing principles and unifying simplifications inherent in special relativity.

The strategy by which we have achieved

$$\text{global gauge} \quad \Longrightarrow \quad \text{local gauge}$$

has "summoned electrodynamics into being," and has at the same time *lent specific structure to the particle–field interaction*. These concluding remarks are intended to expose more clearly some details associated with the latter aspect of our subject. Recall that at (33), working from the ungauged Lagrangian $\mathcal{L}_0$, we extracted definitions

$$\mathcal{E} = R\cdot\tfrac{1}{2m}\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S \equiv RH \qquad \text{and} \qquad \boldsymbol{\mathcal{F}} = \tfrac{1}{m}\mathcal{E}\boldsymbol{\nabla}S = \big(\tfrac{1}{m}R\boldsymbol{\nabla}S\big)H$$

and found

$$\mathcal{E}_t + \boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{F}} = \underbrace{\big\{R_t + \boldsymbol{\nabla}\cdot\big(\tfrac{1}{m}R\boldsymbol{\nabla}S\big)\big\}}_{0} H + \tfrac{1}{m}R\,\underbrace{\big\{\boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}S_t + \boldsymbol{\nabla}S\cdot\boldsymbol{\nabla}H\big\}}_{0}$$
$$= 0 \quad \text{by the ungauged Hamilton-Jacobi equations (22)}$$

To discover the effect of turning on the gauge field we return to (33.10) and (33.20), make the replacement $\mathcal{L}_0 \to \mathcal{L}_1$, and find

$$\left.\begin{aligned} \mathcal{E} &\to \mathcal{E}' = R\cdot H' \\ \boldsymbol{\mathcal{F}} &\to \boldsymbol{\mathcal{F}}' = \big[\tfrac{1}{m}R\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\big]\cdot H' \end{aligned}\right\} \tag{44}$$

with $H' \equiv \frac{1}{2m}\big(\boldsymbol{\nabla}S - \frac{e}{c}\boldsymbol{A}\big)\cdot\big(\boldsymbol{\nabla}S - \frac{e}{c}\boldsymbol{A}\big) + e\phi$.[15] We obtain

$$\begin{aligned} \mathcal{E}'_t + \boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{F}}' &= \underbrace{\big\{R_t + \boldsymbol{\nabla}\cdot\big[\tfrac{1}{m}R\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\big]\big\}}_{0 \quad \text{by (28.2)}} H' + R\big\{H'_t + \big[\tfrac{1}{m}R\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\big]\cdot\boldsymbol{\nabla}H'\big\} \\ &= R\big\{\tfrac{1}{m}\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\cdot\big(\boldsymbol{\nabla}S_t - \tfrac{e}{c}\boldsymbol{A}_t\big) + e\phi_t\big\} - \big[\tfrac{1}{m}R\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\big]\cdot\boldsymbol{\nabla}S_t \\ &= -\tfrac{e}{m}R\big(\boldsymbol{\nabla}S - \tfrac{e}{c}\boldsymbol{A}\big)\cdot\tfrac{1}{c}\boldsymbol{A}_t + eR\phi_t \\ &= \tfrac{1}{c}\,\boldsymbol{J}\cdot\boldsymbol{E} + (\boldsymbol{J}\cdot\boldsymbol{\nabla} + \rho\partial_t)\phi \end{aligned} \tag{45}$$

On the other hand, for the gauge field we found the energy density and flux to be given by

$$\tilde{\mathcal{E}} = \tfrac{1}{2}(\boldsymbol{E}\cdot\boldsymbol{E} + r\boldsymbol{B}\cdot\boldsymbol{B}) + \text{gauge term}$$
$$\tilde{\boldsymbol{\mathcal{F}}} = rc\boldsymbol{E}\times\boldsymbol{B} + \textbf{gauge term}$$

[15] I have found it convenient, for the purposes of this discussion to turn the impressed potential U off; i.e., to assume that our Hamilton-Jacobi theory refers to the motion of particles which—except for gauge field effects—move freely.

It follows readily from the field equations (32) that

$$\tfrac{\partial}{\partial t}\big[\tfrac{1}{2}(\boldsymbol{E}\cdot\boldsymbol{E}+r\boldsymbol{B}\cdot\boldsymbol{B})\big]+\boldsymbol{\nabla}\cdot\big[rc\boldsymbol{E}\times\boldsymbol{B}\big]=-\tfrac{1}{c}\,\boldsymbol{J}\cdot\boldsymbol{E} \tag{46.1}$$

while we established at (37) that

$$\tfrac{\partial}{\partial t}\big[\text{gauge term}\big]+\boldsymbol{\nabla}\cdot\big[\textbf{gauge term}\big]=-(\boldsymbol{J}\cdot\boldsymbol{\nabla}+\rho\partial_t)\phi \tag{46.2}$$

From (45/46) we obtain

$$\tfrac{\partial}{\partial t}\big[\mathcal{E}'+\tilde{\mathcal{E}}\big]+\boldsymbol{\nabla}\cdot\big[\boldsymbol{\mathcal{F}}'+\tilde{\boldsymbol{\mathcal{F}}}\big]=0 \tag{47}$$

which attributes detailed local balance to the energy exchange between the matter field and the gauge field. Three similar results, established by similar means, pertain to local momentum balance.[16]

Note finally that our final (locally gauge-invariant) Lagrangian (28) can be developed (compare (16))

$$\mathcal{L}_{\text{free Hamilton-Jacobi}}+\mathcal{L}_{\text{interaction}}+\mathcal{L}_{\text{free gauge field}}$$

where $\mathcal{L}_{\text{free Hamilton-Jacobi}}$ is just the $\mathcal{L}_0$ which at (23) provided our point of departure, $\mathcal{L}_{\text{free gauge field}}=-\frac{1}{2}(\boldsymbol{E}\cdot\boldsymbol{E}-r\boldsymbol{B}\cdot\boldsymbol{B})$ is (35) with $p=-1$, and, reading from (27),

$$\begin{aligned}\mathcal{L}_{\text{interaction}}&=-\tfrac{1}{c}\tfrac{e}{m}R\big(\boldsymbol{\nabla}S-\tfrac{e}{c}\boldsymbol{A}\big)\cdot\boldsymbol{A}-R\tfrac{1}{2m}\big(\tfrac{e}{c}\big)^2\boldsymbol{A}\cdot\boldsymbol{A}+eR\phi\\&=\rho\phi-\tfrac{1}{c}\big(\boldsymbol{J}+R\tfrac{e^2}{2mc}\boldsymbol{A}\big)\cdot\boldsymbol{A}\\&=\rho\phi-\tfrac{1}{c}\tfrac{1}{2}\big(\boldsymbol{J}+\boldsymbol{j}\big)\cdot\boldsymbol{A}\end{aligned} \tag{48}$$

$$\boldsymbol{j}\equiv\tfrac{e}{m}R\boldsymbol{\nabla}S=\boldsymbol{J}\Big|_{\boldsymbol{A}\to\boldsymbol{0}} \tag{49}$$

Equation (48) mimics the design of an equation to which we were led when looking to the gauge theory of a *relativistic complex* scalar field. The agreement (48) $\leftrightarrow$ (21) becomes in fact precise when one writes[17]

$$\|J^\mu\|\equiv\begin{pmatrix}c\rho\\ \boldsymbol{J}\end{pmatrix}\quad\text{and}\quad\|A^\mu\|=\begin{pmatrix}\phi\\ \boldsymbol{A}\end{pmatrix} \tag{50}$$

and uses the Lorentz metric to lower an index. Equation (48) also conforms to the result achieved when one takes the non-relativistic Schrödinger Lagrangian (1–76) as a point of departure.[18] Pretty clearly: *neither relativity, nor quantum*

[16] Compare (327) on p. 312 of CLASSICAL ELECTRODYNAMICS (1980); the argument there is relativistic, and therefore simpler/briefer.

[17] See CLASSICAL ELECTRODYNAMICS (1980), pp. 161 & 373.

[18] See CLASSICAL FIELD THEORY (1979), 188–193.

mechanics, nor complex-valuedness are essential to the success of the gauge field program. Which was the point at issue—now demonstrated by example.[19]

"Minimal coupling" and the physical significance of current. We consider those subjects in reverse order, taking as our point of departure this question: How does electrical current—so "real" it can be measured with an ammeter—come to be represented in theory by an expression which is (on its face) not even gauge-invariant? To expose the points at issue in their simplest and most essential terms I look to the relativistic classical mechanics of a particle.

Let $x^\mu(\tau)$, $u^\mu(\tau) \equiv \frac{d}{d\tau}x^\mu(\tau)$ and $a^\mu(\tau) \equiv \frac{d}{d\tau}u^\mu(\tau)$ refer the position (with respect to an intertial frame), 4-velocity and 4-acceleration of a mass point m. From the definition of proper time τ it follows that $(u,u) \equiv g_{\alpha\beta}u^\alpha u^\beta = c^2$, and therefore that $(a,u) = 0$; i.e., that $a \perp u$ in the Lorentzian sense. Minkowski's equation of motion reads $ma^\mu = K^\mu$. Necessarily, $(K,u) = 0$: Minkowski forces K^μ are necessarily velocity-dependent. In the simplest case K^μ will depend *linearly* upon 4-velocity: $K^\mu \sim F^{\mu\nu}u_\nu$. From $K^\mu u_\mu = 0$ (all u) it follows that necessarily $F^{\mu\nu}$ is *antisymmetric*. We are led thus to consider relativistic systems of the especially simple design

$$ma^\mu = \tfrac{e}{c}F^{\mu\nu}(x)\,u_\nu \tag{51}$$

where $[eF^{\mu\nu}]$ = (force) and e is a coupling constant. Passing now from the Minkowskian to the Lagrangian side of the street[20]...

The simplest way to build velocity-dependence into a Lagrangian is to write

$$L = \tfrac{1}{2}mg_{\alpha\beta}u^\alpha u^\beta + \tfrac{e}{c}A_\alpha(x)u^\alpha \tag{52}$$

[19] I find the example to be of some intrinsic interest. The possibility of a "gauged Hamilton-Jacobi theory" has been known to me since the spring of 1984, when I wrote the material which appears on pp. 489–502 of CLASSICAL MECHANICS (1983–1984), but I had forgotten until this minute that I had ever actually written about the subject; my field theory books provide only the rough seminar notes presented under the title "A sense in which classical mechanics quantizes itself" (1980) and bound as an appendix to CLASSICAL FIELD THEORY (1979). Freshly emerged now from such a writing experience, I am filled with a renewed sense of *what a wonderous activity-organizer is the Lagrangian formalism!* It asks you to do some occasionally intricate things... which, however, seem preordained always to work out. I am impressed also by how awkward life can become when relativity is removed from one's toolbox.

[20] This is less easily accomplished than might be supposed, for the constraint $(u,u) = c^2$ must be folded into the meaning of the variational process $\delta\int L\,d\tau$. Various techniques for accomplishing that objective are described in the iintroductory sections of RELATIVISTIC DYNAMICS (1975). But the following discussion leads to equations of motion for which compliance with the constraint becomes—as it happens—*automatic*, and it is this lucky circumstance which permits us to set such subtleties aside.

We are led then to equations of the motion of the form

$$\Big\{\frac{d}{d\tau}\frac{\partial}{\partial u^\mu} - \frac{\partial}{\partial x^\mu}\Big\}L = ma_\mu + \tfrac{e}{c}A_{\mu,\alpha}(x)u^\alpha - \tfrac{e}{c}A_{\alpha,\mu}(x)u^\alpha = 0$$

which can be written

$$ma^\mu = \tfrac{e}{c}F^{\mu\nu}(x)\,u_\nu \quad \text{with} \quad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{53}$$

These equations—which we may interpret to refer to the relativistic motion of a charged mass point in the presence of an impressed electromagnetic field—are invariant under

$$A_\mu \longrightarrow A_\mu + \partial_\mu\Omega \tag{54.1}$$

but derive from a Lagrangian which is *not* gauge-invariant:

$$\begin{aligned} L &\longrightarrow L + \text{offending term} \\ &\text{offending term} = \tfrac{e}{c}u^\alpha\partial_\alpha\Omega = \tfrac{d}{d\tau}\big[\tfrac{e}{c}\,\Omega(x)\big] \end{aligned}$$

Notice, however, that we have only to assign an expanded meaning

$$A_\mu \longrightarrow A_\mu + \partial_\mu\Omega \qquad \rightsquigarrow \qquad \begin{cases} A_\mu \longrightarrow A_\mu + \partial_\mu\Omega(x) \\ L \longrightarrow L - \frac{d}{d\tau}\big[\frac{e}{c}\,\Omega(x)\big] \end{cases} \tag{54.2}$$

to the notion of a "gauge transformation" to acquire gauge-invariance of the Lagrangian, whence of all that follows from the Lagrangian. To illustrate the point:

Working from (52), we find the momentum conjugate to x^μ to be given by

$$p_\mu = \frac{\partial L}{\partial u^\mu} = mu_\mu + \tfrac{e}{c}A_\mu \tag{55}$$

which

- is not gauge-invariant under the interpretation (54.1), but
- is gauge-invariant under the expanded interpretation (54.2).

To say the same thing another way: gauge transformations, under the restricted interpretation (54.1), send

$$\left.\begin{aligned} L &\longrightarrow L + \tfrac{e}{c}u^\alpha\partial_\alpha\Omega \\ p_\mu &\longrightarrow p_\mu + \tfrac{e}{c}\partial_\mu\Omega \end{aligned}\right\} \tag{56.1}$$

but under the expanded interpretation (54.2) send

$$\left.\begin{aligned} L &\longrightarrow \big\{L - \tfrac{d}{d\tau}\big[\tfrac{e}{c}\,\Omega(x)\big]\big\} + \tfrac{e}{c}u^\alpha\partial_\alpha\Omega = L \\ p_\mu &\longrightarrow \big\{p_\mu - \tfrac{\partial}{\partial u^\mu}\tfrac{d}{d\tau}\big[\tfrac{e}{c}\,\Omega(x)\big]\big\} + \tfrac{e}{c}\partial_\mu\Omega = p_\mu \end{aligned}\right\} \tag{56.2}$$

Passing now to the Hamiltonian formalism, we find that

$$H(p,x) = p_\alpha u^\alpha - L(x,u) \qquad \text{with} \qquad u \mapsto \tfrac{1}{m}\big[p - \tfrac{e}{c}A\big] \tag{57}$$

gives

$$= \tfrac{1}{2m} g^{\alpha\beta}\big[p_\alpha - \tfrac{e}{c}A_\alpha\big]\big[p_\beta - \tfrac{e}{c}A_\beta\big] \tag{58}$$

The resulting canonical equations

$$\begin{aligned} u^\mu &= +\tfrac{\partial H}{\partial p_\mu} = \tfrac{1}{m} g^{\mu\alpha}\big[p_\alpha - \tfrac{e}{c}A_\alpha\big] \\ \tfrac{d}{d\tau}p_\mu &= -\tfrac{\partial H}{\partial x^\mu} = \tfrac{1}{m} g^{\alpha\beta}\big[p_\alpha - \tfrac{e}{c}A_\alpha\big]\tfrac{e}{c}A_{\beta,\mu} \end{aligned}$$

are readily seen to reproduce (53).[21] The gauge transformation properties of the Hamiltonian are somewhat subtle, and to sort them out I need to distinguish "L-gauge" ($L \longrightarrow L + u^\alpha \partial_\alpha \Lambda(x)$) from "$A$-gauge" ($A_\mu \longrightarrow A_\mu u + \partial_\alpha \Omega(x)$). The former causes the $H(p,x)$ of (57) to go over into

$$[p_\alpha + \partial_\alpha\Lambda]u^\alpha - [L + u^\alpha\partial_\alpha\Lambda] \qquad \text{with} \qquad u \mapsto \tfrac{1}{m}\big([p - \partial\Lambda] - \tfrac{e}{c}A\big)$$

In short (note the cancellation): L-gauge causes

$$H(p,x) \longrightarrow H(p - \partial\Lambda, x)$$

while reading from (58) we see that that A-gauge causes

$$H(p,x) \longrightarrow H(p - \tfrac{e}{c}\partial\Omega, x)$$

These elementary remarks expose in new light the central idea of gauge field theory: *use one gauge type to cancel the effect of the other*, by setting $\Lambda = -\frac{e}{c}\Omega$.

Electrical "current" enters the discussion as a by-product of steps we take in order to promote the A_μ-field—heretofore considered to have been externally impressed/prescribed—to the status of a *dynamical field in its own right*. From the A-gauge-invariant antisymmetric tensor field $F^{\mu\nu}$ on can—using $g_{\mu\nu}$ and $\epsilon_{\mu\nu\rho\sigma}$ as "glue"—construct a total of three Lorentz invariants:

$$F^\mu{}_\nu F^\nu{}_\mu\,, \quad F^\mu{}_\nu G^\nu{}_\mu\,, \quad \text{and} \quad G^\mu{}_\nu G^\nu{}_\mu \quad \text{with} \quad G_{\mu\nu} \equiv \tfrac{1}{2}\epsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$$

which in index-free notation can be described

$$\mathrm{tr}\,\mathbb{F}\mathbb{F}, \quad \mathrm{tr}\,\mathbb{F}\mathbb{G}, \quad \text{and} \quad \mathrm{tr}\,\mathbb{G}\mathbb{G}$$

[21] ...and can be considered to arise as "meta-Lagrange equations"

$$\Big\{\frac{d}{d\tau}\frac{\partial}{\partial \dot{p}_\mu} - \frac{\partial}{\partial p_\mu}\Big\}\mathfrak{L} = 0 \quad \text{and} \quad \Big\{\frac{d}{d\tau}\frac{\partial}{\partial u^\mu} - \frac{\partial}{\partial x^\mu}\Big\}\mathfrak{L} = 0$$

from the "meta-Lagrangian"

$$\mathfrak{L} \equiv p_\alpha u^\alpha - H(p,x)$$

But it is not difficult to show[22] that

$$\mathrm{tr}\mathbb{G}\mathbb{G} \sim \mathrm{tr}\mathbb{F}\mathbb{F} \quad \text{and} \quad \mathrm{tr}\mathbb{F}\mathbb{G} \sim \partial_\mu\big[\epsilon^{\mu\nu\rho\sigma}A_\nu(\partial_\rho A_\sigma - \partial_\sigma A_\rho)\big]$$

So $\mathrm{tr}\mathbb{F}\mathbb{F}$ and $\mathrm{tr}\mathbb{G}\mathbb{G}$ contribute identically (apart from a trivial factor) when introduced into a Lagrangian, while $\mathrm{tr}\mathbb{F}\mathbb{G}$ contributes only an inconsequential gauge term. We are led, therefore, to examine

$$\begin{aligned} L_2 &= \tfrac{1}{2}mg_{\alpha\beta}u^\alpha u^\beta + \tfrac{e}{c}A_\alpha(x)u^\alpha + \tfrac{1}{4}F_{\alpha\beta}F^{\alpha\beta} \\ &= \tfrac{1}{2}mg_{\alpha\beta}u^\alpha u^\beta + \tfrac{e}{c}A_\alpha(x)u^\alpha + \tfrac{1}{2}(g^{\alpha\rho}g^{\beta\sigma} - g^{\alpha\sigma}g^{\beta\rho})A_{\alpha,\beta}A_{\rho,\sigma} \end{aligned} \tag{59}$$

From

$$\Big\{\partial_\mu\frac{\partial}{\partial A_{\nu,\mu}} - \frac{\partial}{\partial A_\nu}\Big\}L_2 = \partial_\mu(A^{\nu,\mu} - A^{\mu,\nu}) - \frac{\partial L}{\partial A_\nu} = 0$$

we obtain

$$\partial_\mu F^{\mu\nu} = \tfrac{1}{c}J^\nu \tag{60}$$

with

$$J^\nu \equiv \frac{\partial L}{\partial A_\nu} = eu^\nu \tag{61.1}$$

$$= \tfrac{e}{m}\big[p^\nu - \tfrac{e}{c}A^\nu\big] \tag{61.2}$$

The expression on the right in (61.1) is—by every interpretation—manifestly gauge-invariant, and conforms precisely to what, on physical grounds, we expect of the "4-current of a charged mass point." The expression of the right in (6.12) is, on the other hand, gauge-invariant only under the expanded interpretation (54.2), and it is under the latter interpretation that the Lagrangian of (59) becomes gauge-invariant. The notation (61.1) permits the interaction term present in (59) to be described

$$\mathcal{L}_{\text{interaction}} = \tfrac{1}{c}J^\alpha A_\alpha(x) \tag{62}$$

This is the term which appears, on its face, to mess up gauge-invariance, but which becomes gauge-invariant in the expanded sense; it was precisely the gauge-failure of $\mathcal{L}_{\text{interaction}}$ that the L-gauge of the overall Lagrangian was tailored to correct.

Passing from (59) to the equivalent Hamiltonian formalism, one has

$$\begin{aligned} H(p,x) &= \tfrac{1}{2m}g^{\alpha\beta}\big[p_\alpha - \tfrac{e}{c}A_\alpha\big]\big[p_\beta - \tfrac{e}{c}A_\beta\big] - \tfrac{1}{4}F_{\alpha\beta}F^{\alpha\beta} \\ &= \tfrac{1}{2m}g^{\alpha\beta}\big[p_\alpha - \tfrac{e}{c}A_\alpha\big]\big[p_\beta - \tfrac{e}{c}A_\beta\big] - \tfrac{1}{2}(g^{\alpha\rho}g^{\beta\sigma} - g^{\alpha\sigma}g^{\beta\rho})A_{\alpha,\beta}A_{\rho,\sigma} \end{aligned} \tag{63}$$

In this formalism the extended gauge transformation (54.2) lurks behind the scenery, and presents this face:

$$\left.\begin{aligned} A_\mu &\longrightarrow A_\mu + \partial_\mu\Omega(x) \\ p_\mu &\longrightarrow p_\mu - \tfrac{e}{c}\partial_\mu\Omega(x) \end{aligned}\right\} \tag{64}$$

[22] See CLASSICAL ELECTRODYNAMICS (1980), p. 298.

The invariance of (63) under (64) is manifest. As was noted already in the discussion subsequent to (58), the canonical equations implicit in (63) reproduce our initial description (53) of the dynamics of the charged particle. But what of the dynamics of the gauge field? Yielding unthinkingly to entrenched habit, we construct

$$\Big\{\partial_\mu \frac{\partial}{\partial A_{\nu,\mu}} - \frac{\partial}{\partial A_\nu}\Big\} H = -\partial_\mu (A^{\nu,\mu} - A^{\mu,\nu}) - \frac{\partial H}{\partial A_\nu} = 0$$

which does indeed give back (60/61):

$$\partial_\mu F^{\mu\nu} = -\frac{\partial H}{\partial A_\nu} = \tfrac{e}{mc}\big[p^\nu - \tfrac{e}{c}A^\nu\big] = \tfrac{1}{c}J^\nu$$

I say "unthinkingly" because we have no secure reason to take the Lagrange derivative of a Hamiltonian! It would, I think, be better form to construct the "meta-Lagrangian"

$$\mathfrak{L}(p, x, \bullet, u, A, \partial A) = p_\alpha u^\alpha - \big\{\tfrac{1}{2m} g^{\alpha\beta}\big[p_\alpha - \tfrac{e}{c}A_\alpha\big]\big[p_\beta - \tfrac{e}{c}A_\beta\big] - \tfrac{1}{4}F_{\alpha\beta}F^{\alpha\beta}\big\}$$

and recover the canonical equations as "meta-Lagrange equations," but we would come out in the same place.[21]

The Hamiltonian (63) can be developed

$$\begin{aligned} H = \tfrac{1}{2m} g^{\alpha\beta} p_\alpha p_\beta - H_{\text{int}} - \tfrac{1}{4}F_{\alpha\beta}F^{\alpha\beta} & \\ H_{\text{int}} \equiv \tfrac{e}{c}\tfrac{1}{m}\big[p^\alpha - \tfrac{e}{c}A^\alpha\big]A_\alpha + \tfrac{1}{2m}\big(\tfrac{e}{c}\big)^2 A^\alpha A_\alpha & \\ = \tfrac{1}{c}J^\alpha A_\alpha + \tfrac{e^2}{2mc^2}A^\alpha A_\alpha & \qquad (65) \end{aligned}$$

though to do so entails a term-by-term sacrifice of manifest gauge-invariance. If we borrow notation from (18.1), writing $J^\nu \equiv j^\nu - \frac{e^2}{mc}A^\nu$, then we have

$$\begin{aligned} = \tfrac{1}{c}\tfrac{1}{2}(J^\alpha + j^\alpha)A_\alpha & \qquad (66) \\ j^\alpha \equiv \tfrac{e}{m}p^\alpha = J^\alpha\Big|_{A\to 0} & \end{aligned}$$

which mimics (21) and (48). Notice, however, this curious circumstance: the second term on the right side of (65)—which at (21) read $\frac{e^2}{2mc^2}(\psi^* A^\alpha \psi)A_\alpha$, and at (48) read $\frac{e^2}{2mc^2}R\,\boldsymbol{A}\cdot\boldsymbol{A}$—displays now no reference to the particle; only the e^2 reveals the "interactive" nature of the term, which we might otherwise be tempted to classify as a "mass term" present in the design of $H_{\text{free gauge field}}$.

What have we learned?

People sometimes point to (62)—i.e., to terms of the design $\boldsymbol{J}\cdot\boldsymbol{A}$, which in Lagrangian formalism serve to describe the interaction of charged matter with the electromagnetic field—as the defining symptom of "minimal coupling." But more standardly, the term is taken[23] to refer to the characteristic matter-field

[23] See, for example, M. E. Peskin & D. V. Schroeder, *An Introduction to Quantum Field Theory* (1995), p. 78.

interaction which arises from pursuit of the gauge field program; i.e., which springs spontaneously from $p \longrightarrow p - (e/c)A$ (or again: from $\partial \longrightarrow \mathcal{D}$).

Gauge field theory usually has a quantum mechanical objective, and for that reason is strongly Hamiltonian in spirit, though presented as an exercise in Lagrangian field theory.[24] In the examples we have studied we have been led at (21/48/66) to interaction terms which are "of a type," but more complicated than is suggested by the more purely Lagrangian model (62). We have learned, however, that the isolation of "interaction terms" $\mathcal{L}_{\text{int}}$ typically violates the "principle of manifest gauge-invariance," and thus runs counter to the essential spirit of gauge field theory; it is better to allow the interaction to remain implicit, as (for example) it did when at (19) and (17) we wrote

$$\begin{aligned} g^{\alpha\beta}\big(\tfrac{\hbar}{i}\partial_\alpha - \tfrac{e}{c}A_\alpha\big)\big(\tfrac{\hbar}{i}\partial_\beta - \tfrac{e}{c}A_\beta\big)\psi &= (mc)^2\psi \\ \partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) &= \tfrac{1}{c}J^\nu \end{aligned}$$

with $J^\nu = e\big[-i\tfrac{\hbar}{2m}g^{\nu\alpha}\big(\psi^*_{,\alpha}\psi - \psi^*\psi_{,\alpha}\big) + \tfrac{e}{mc}\big(\psi^*\psi\big)A^\nu\big]$.

Gauge field theory standardly takes a "field theory of matter" as its point of departure, but we have learned that the gauge field concept is so primitive that one can *abandon the initial field theory*: we achieved success when we proceeded from the Lagrangian mechanics of a single relativistic particle... and might (with some loss of simplicity) have abandoned the relativity; the resulting theory would have captured the simple essence of our "gauged Hamilton-Jacobi theory."

We have learned that the gauge-invariance of the currents that arise from gauge field theory is invariably *present but covert*. Our particulate model supplied

$$J^\mu = \frac{e}{m}\big[p^\mu - \frac{e}{c}A^\mu\big] = eu^\mu$$

which suggests that the complexity of the expressions that serve, in various contexts, to define J^μ can be attributed to the familiar complexity of the relationship between "velocity" (a physical observable, at least in particle mechanics) and "conjugate momentum" (a theoretical construct).

In the beginning was a decision—a decision to "allow the phase factor vary from point to point"—which may at the time have seemed willful, arbitrary, justifiable only by the illuminating results to which it could be shown to lead. Our particulate model has allowed that decision to be replaced by a formal act which I find much more natural: require of the formulæ of (meta-)Lagrangian dynamics that they be manifestly invariant under arbitrary gauge-adjustments

$$L \longrightarrow L + \partial_\alpha \Lambda^\alpha$$

I shall on another occasion describe how gauge field theory might procede from such a starting point (that effort will require only rearrangement of what I have

[24] Both remarks, by the way, pertain also to our "gauged Hamilton-Jacobi theory."

already written, and some shifted emphasis), but turn now to more pressing matters.

Gauged Dirac theory. We take now as our point of departure the Lagrangian

$$\mathcal{L}_0(\psi,\tilde{\psi},\partial\psi,\partial\tilde{\psi}) = \hbar c\Big[\tfrac{1}{2}i\big\{\tilde{\psi}\boldsymbol{\gamma}^\alpha\psi_{,\alpha} - \tilde{\psi}_{,\alpha}\boldsymbol{\gamma}^\alpha\psi\big\} - \varkappa\tilde{\psi}\psi\Big] \tag{67}$$

which was seen at (2–56) to yield the Dirac equations

$$\big(\boldsymbol{\gamma}^\mu\partial_\mu + i\varkappa\big)\psi = 0 \quad \text{and its adjoint} \tag{68}$$

From the manifest *global* phase-invariance of $\mathcal{L}_0$—infinitesimally: from the invariance of $\mathcal{L}_0$ under

$$\begin{aligned}\psi &\longrightarrow \psi + \delta_\omega\psi \quad \text{with} \quad \delta_\omega\psi = +i\psi\cdot\delta\omega\\ \tilde{\psi} &\longrightarrow \tilde{\psi} + \delta_\omega\tilde{\psi} \quad \text{with} \quad \delta_\omega\tilde{\psi} = -i\tilde{\psi}\cdot\delta\omega\end{aligned}$$

—we obtain the conservation law

$$\partial_\mu Q^\mu = 0 \tag{69.1}$$

$$Q^\mu \equiv -\tfrac{1}{\hbar}\left\{\frac{\partial\mathcal{L}_0}{\partial\psi_{,\mu}}(i\psi) + \frac{\partial\mathcal{L}_0}{\partial\tilde{\psi}_{,\mu}}(-i\tilde{\psi}\,)\right\} = c\,\tilde{\psi}\boldsymbol{\gamma}^\mu\psi \tag{69.2}$$

where an $\hbar$-factor has been introduced so as to achieve[25]

$$[Q^\mu] = \frac{1}{\text{area}\cdot\text{time}} = \text{number flux}$$

Our objective is to achieve *local* phase invariance. Familiar steps lead us, therefore, to construction of the system

$$\begin{aligned}&\mathcal{L}_2(\psi,\tilde{\psi},\partial\psi,\partial\tilde{\psi},A,\partial A) = \mathcal{L}_0(\psi,\tilde{\psi},[\partial - i\tfrac{e}{\hbar c}A]\psi,[\partial + i\tfrac{e}{\hbar c}A]\tilde{\psi}) + \tfrac{1}{4}F^{\alpha\beta}F_{\alpha\beta}\\ &= \hbar c\Big[\tfrac{1}{2}i\Big\{\tilde{\psi}\boldsymbol{\gamma}^\alpha(\psi_{,\alpha} - i\tfrac{e}{\hbar c}A_\alpha\psi) - (\tilde{\psi}_{,\alpha} + i\tfrac{e}{\hbar c}A_\alpha\tilde{\psi})\boldsymbol{\gamma}^\alpha\psi\Big\} - \varkappa\tilde{\psi}\psi\Big] + \tfrac{1}{4}F^{\alpha\beta}F_{\alpha\beta}\\ &= \mathcal{L}_0(\psi,\tilde{\psi},\partial\psi,\partial\tilde{\psi}) + \tfrac{1}{C}J^\alpha A_\alpha + \tfrac{1}{4}F^{\alpha\beta}F_{\alpha\beta}\end{aligned} \tag{70}$$

where the gauge-invariant antisymmetric tensor field $F_{\mu\nu}$ retains its former definition (15) and where

$$J^\mu \equiv eQ^\mu \tag{71}$$

[25] Notice now much simpler (69.2) is than were its non-relativistic/relativistic scalar counterparts (1–87) and (3–13). That simplicity can be attributed to the circumstance that $\mathcal{L}_0$ is *linear in the derivatives* of the complex field. It is, therefore, a simplicity not special to the Dirac theory, but shared by all canonically formulated theories... of which the Dirac theory provides merely the simplest instance.

Only once before—at (61.1) and (62), in connection with our particulate model —have we encountered formulæ so simple as those which serve here to describe J^μ and $\mathcal{L}_{\text{interaction}}$; we note in particular that the distinction between J^μ and j^μ does not force itself upon our attention in Dirac theory.

The field equations which arise from the twice-modified (i.e., from the "gauged and launched") Lagrangian (70) read

$$\left.\begin{array}{c} \left[\boldsymbol{\gamma}^\mu(\partial_\mu - i\frac{e}{\hbar c}A_\mu) + i\varkappa\right]\psi = 0 \quad \text{and its adjoint} \\ \partial_\mu F^{\mu\nu} = \frac{1}{C}J^\nu \end{array}\right\} \tag{72.1}$$

with

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu \tag{72.2}$$

The "synchronized local gauge transformation" with respect to which the theory is—by design—invariant can in present notation be described

$$\left.\begin{array}{l} \psi \longrightarrow \psi' = e^{ig\Omega(x)} \cdot \psi \\ A_\mu \longrightarrow A'_\mu = A_\mu + \partial_\mu \Omega(x) \end{array}\right\} \tag{73}$$

with $g = e/\hbar c$. Those statements were contrived to entail

$$(\partial_\mu - igA_\mu)\psi \longrightarrow e^{ig\Omega(x)} \cdot (\partial_\mu - igA_\mu)\psi \tag{74}$$

and it is, to reiterate, that contrivance—first encountered at (9)—which lies at the heart of guage field theory.

Mathematical interlude: non-Abelian gauge groups. The operations "multiply by a phase factor"—which when ψ is an N-component complex field have this explicit meaning:

$$\begin{pmatrix} \psi^1 \\ \psi^2 \\ \vdots \\ \psi^N \end{pmatrix} \longrightarrow \begin{pmatrix} e^{i\omega} & 0 & \dots & 0 \\ 0 & e^{i\omega} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\omega} \end{pmatrix} \begin{pmatrix} \psi^1 \\ \psi^2 \\ \vdots \\ \psi^N \end{pmatrix}$$

—clearly possess the group property. The abstract group in question is $U(1)$, of which the 1×1 unitary matrices $\boldsymbol{U}(\omega) \equiv \|e^{i\omega}\| = e^{i\omega}\boldsymbol{I}$ provide the simplest representation. It becomes natural in this light to write

$$\begin{pmatrix} \psi^1 \\ \psi^2 \\ \vdots \\ \psi^N \end{pmatrix} \longrightarrow \begin{pmatrix} U^1{}_1 & U^1{}_2 & \dots & U^1{}_N \\ U^2{}_1 & U^2{}_2 & \dots & U^2{}_N \\ \vdots & \vdots & \ddots & \vdots \\ U^N{}_1 & U^N{}_2 & \dots & U^N{}_N \end{pmatrix} \begin{pmatrix} \psi^1 \\ \psi^2 \\ \vdots \\ \psi^N \end{pmatrix} \equiv \boldsymbol{U} \begin{pmatrix} \psi^1 \\ \psi^2 \\ \vdots \\ \psi^N \end{pmatrix}$$

and to admit the possibility that $\boldsymbol{U}$ might be an element of the group $U(N)$ of $N \times N$ unitary matrices. This is the idea which Yang & Mills (1954) were,

for their own good reasons,[26] the first to explore... with results which first alerted physicists to the possibility that gauge field theory might be put to more informative uses that the "elegant re-invention of electrodynamics." My objective here will to assemble the mathematical material we will need to pursue that idea.

From the unitarity of $\boldsymbol{U}$ it follows that $(\det \boldsymbol{U})^*(\det \boldsymbol{U}) = 1$, and therefore that

$$\det \boldsymbol{U} = e^{i\vartheta}$$

Write $\boldsymbol{U} = e^{i\boldsymbol{H}}$ and observe that $\boldsymbol{U}$ will be unitarity $\boldsymbol{U}^\dagger = \boldsymbol{U}^{-1}$ if and only if $\boldsymbol{H}$ is hermitian. A general identity supplies $\det \boldsymbol{U} = \exp\big\{i(\mathrm{tr}\boldsymbol{H})\big\}$, from which we infer that $\vartheta = \mathrm{tr}\boldsymbol{H}$. If $\boldsymbol{U}$ is unitary then so is $\boldsymbol{S} \equiv e^{-i\omega}\boldsymbol{U}$, and $\det \boldsymbol{S} = e^{i(\vartheta - N\omega)}$, where N refers to the *dimension* of $\boldsymbol{U}$. We have only to set $\omega = \vartheta/N$ to render $\boldsymbol{S}$ *unimodular*; i.e., to achieve $\det \boldsymbol{S} = 1$. Unimodularity is preserved under multiplication: the $N \times N$ matrices $\boldsymbol{S}$ are elements of a subgroup (denoted $SU(N)$ and called the "special unitary group") of $U(N)$. One writes

$$U(N) = U(1) \otimes SU(N)$$

to signify that every element of $U(N)$ can be written

$$\boldsymbol{U} = e^{i\omega} \cdot \boldsymbol{S} \quad \text{with} \quad \begin{cases} e^{i\omega} \in U(N) \\ \boldsymbol{S} \in SU(N) \end{cases}$$

Matrices $\boldsymbol{S} \in SU(N)$ can be written

$$\boldsymbol{S} = e^{i\boldsymbol{H}} \quad \text{where } \boldsymbol{H} \text{ is a } \textit{traceless} \text{ hermitian matrix}$$

The most general such matrix $\boldsymbol{H}$ can be displayed

$$\boldsymbol{H} = \begin{pmatrix} d_1 & a_1 + ib_1 & a_2 + ib_2 & \dots \\ a_1 - ib_1 & d_2 & a_N + ib_N & \dots \\ a_2 - ib_2 & a_N - ib_N & d_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad \text{with } d_1 + d_2 + \cdots + d_N = 0$$

and contains $N^2 - 1$ adjustable constants. The set of such matrices is closed under addition and multiplication by real numbers, so can be considered to comprise a real vector space $V_{\mathcal{N}}$ of $\mathcal{N} \equiv N^2 - 1$ dimensions. Select any basis $\big\{\boldsymbol{h}_1, \boldsymbol{h}_2, \dots, \boldsymbol{h}_{\mathcal{N}}\big\}$ in $V_{\mathcal{N}}$. The general element of $V_{\mathcal{N}}$ can then be developed

$$\boldsymbol{H} = H^1\boldsymbol{h}_1 + H^2\boldsymbol{h}_2 + \cdots + H^{\mathcal{N}}\boldsymbol{h}_{\mathcal{N}}$$

Bases can, of course, be selected in limitlessly many ways, but some algebraic "principles of selection" will soon emerge.

---

[26] Those reasons are evident already in the title of their paper: "Conservation of isotopic spin and isotopic gauge invariance," Phys. Rev. **96**, 191 (1954).

Look to the case $N = 2$. Pauli (see again (2–54)) would in that case have us write

$$\boldsymbol{H} = H^1\boldsymbol{\sigma}_1 + H^2\boldsymbol{\sigma}_2 + H^3\boldsymbol{\sigma}_3 = \begin{pmatrix} H^3 & H^1 - iH^2 \\ H^1 + iH^2 & -H^3 \end{pmatrix} \tag{75}$$

for the reason that the $\boldsymbol{\sigma}$-matrices thus defined are endowed with some especially attractive/useful algebraic properties:

$$\boldsymbol{\sigma}_1\boldsymbol{\sigma}_1 = \boldsymbol{\sigma}_2\boldsymbol{\sigma}_2 = \boldsymbol{\sigma}_3\boldsymbol{\sigma}_3 = \bar{\boldsymbol{I}} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{76.1}$$

$$\left.\begin{aligned} \boldsymbol{\sigma}_1\boldsymbol{\sigma}_2 &= i\,\boldsymbol{\sigma}_3 = -\boldsymbol{\sigma}_2\boldsymbol{\sigma}_1 \\ \boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 &= i\,\boldsymbol{\sigma}_1 = -\boldsymbol{\sigma}_3\boldsymbol{\sigma}_2 \\ \boldsymbol{\sigma}_3\boldsymbol{\sigma}_1 &= i\,\boldsymbol{\sigma}_2 = -\boldsymbol{\sigma}_1\boldsymbol{\sigma}_3 \end{aligned}\right\} \tag{76.2}$$

$$\left.\begin{aligned} [\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2] &= 2i\,\boldsymbol{\sigma}_3 \\ [\boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3] &= 2i\,\boldsymbol{\sigma}_1 \\ [\boldsymbol{\sigma}_3, \boldsymbol{\sigma}_1] &= 2i\,\boldsymbol{\sigma}_2 \end{aligned}\right\} \tag{76.3}$$

Look similarly to the case $N = 3$. Gell-Mann[27] found it convenient to write

$$\begin{aligned} \boldsymbol{H} &= H^1\boldsymbol{\lambda}_1 + H^2\boldsymbol{\lambda}_2 + H^3\boldsymbol{\lambda}_3 + H^4\boldsymbol{\lambda}_4 + H^5\boldsymbol{\lambda}_5 + H^6\boldsymbol{\lambda}_6 + H^7\boldsymbol{\lambda}_7 + H^8\boldsymbol{\lambda}_8 \\ &= \begin{pmatrix} H^3 + \frac{1}{\sqrt{3}}H^8 & H^1 - iH^2 & H^4 - iH^5 \\ H^1 + iH^2 & -H^3 + \frac{1}{\sqrt{3}}H^8 & H^6 - iH^7 \\ H^4 + iH^5 & H^6 + iH^7 & -\frac{2}{\sqrt{3}}H^8 \end{pmatrix} \end{aligned} \tag{77}$$

with algebraic consequences so relatively complicated that I will postpone discussion of them until it has become clearer what it is that we want to know.[28]

Our interest in the *multiplicative*—as opposed to the merely additive—properties of the basic matrices $\boldsymbol{h}_a$ comes to the fore when they are pressed into service as the "generators" of finite transformations, in the sense

$$\boldsymbol{S} \equiv e^{i\,\boldsymbol{H}} = \lim_{m\to\infty}\left[\bar{\boldsymbol{I}} + \tfrac{1}{m}(i\boldsymbol{H})\right]^m$$

Certainly we would develop a lively interest in algebra if we set out to obtain a closed-form *evaluation* of $\sum \frac{1}{n!}(i\boldsymbol{H})^n$. But it is from another (simpler) quarter that our algebraic interest actually springs:

---

[27] M. Gell-Mann & Y. Ne'eman, *The Eightfold Way* (1964), p. 49; see also p. 502 in Peskin & Schroeder,[23] and CLASSICAL GAUGE FIELDS (1981), p. 44.

[28] In the meantime, see Appendix A.3 in M. Kaku, *Quantum Field Theory: A Modern Introduction* (1993).

Generally, the elements of $SU(N)$ fail to commute (which is all one means when one says of $SU(N)$ that it is—in contrast to the "Abelian" group $U(1)$ —"non-Abelian"): $\boldsymbol{S}_1\boldsymbol{S}_2 \neq \boldsymbol{S}_2\boldsymbol{S}_1$. One has[29]

$$\begin{aligned}\boldsymbol{S}_1\boldsymbol{S}_2 &= e^{i\boldsymbol{H}_1}e^{i\boldsymbol{H}_2}\\ &= e^{i(\boldsymbol{H}_1+\boldsymbol{H}_2)-\frac{1}{2}[\boldsymbol{H}_1,\boldsymbol{H}_2]+\text{higher order nested commutators}}\\ &= \boldsymbol{S}_2\boldsymbol{S}_1 \quad \text{if and only if } [\boldsymbol{H}_1,\boldsymbol{H}_2]=\boldsymbol{0}\end{aligned}$$

It was Sophus Lie who first appreciated that the "group multiplication table" (which would supply the evaluation of $\boldsymbol{S}_1\boldsymbol{S}_2$ in all instances) is latent in the commutation properties of the generators. And, moreover, that one can in all cases expect to obtain relations of the form

$$[\text{generator}, \text{generator}] = \text{linear combination of generators} \tag{78}$$

which stands as the defining characteristic of the theory of *Lie algebras*. In the present context (78) becomes

$$[\boldsymbol{h}_p, \boldsymbol{h}_q] = i\sum_{r=1}^{\mathcal{N}} c_p{}^r{}_q \boldsymbol{h}_r \tag{79}$$

The i reflects the elementary circumstance that

$$\begin{aligned}[\text{hermitian}, \text{hermitian}] &= \text{traceless antihermitian}\\ &= i\,(\text{traceless hermitian})\end{aligned}$$

and the real numbers $c_p{}^r{}_q$ are the *structure constants* characteristic of the group. At (76.3) we have already encountered a particular instance of (79).

The structure constants are not freely assignable, but subject to certain constraints. From the antisymmetry of the commutator it follows, for example, that

$$c_p{}^r{}_q = -c_q{}^r{}_p \tag{80.1}$$

while from Jacobi's identity, written $[\boldsymbol{h}_p,[\boldsymbol{h}_q,\boldsymbol{h}_n]]-[\boldsymbol{h}_q,[\boldsymbol{h}_p,\boldsymbol{h}_n]]=[[\boldsymbol{h}_p,\boldsymbol{h}_q],\boldsymbol{h}_n]$, we obtain

$$c_p{}^m{}_k\, c_q{}^k{}_n - c_q{}^m{}_k\, c_p{}^k{}_n = c_p{}^r{}_q \cdot c_r{}^m{}_n \tag{80.2}$$

which can be written

$$\mathbb{C}_p\mathbb{C}_q - \mathbb{C}_q\mathbb{C}_p = i\,c_p{}^r{}_q\,\mathbb{C}_r \tag{81}$$

Evidently the imaginary $\mathcal{N}\times\mathcal{N}$ matrices $\mathbb{C}_r \equiv \|i\,c_r{}^m{}_n\|$ $(r=1,2,\ldots,\mathcal{N})$ provide a representation (the so-called "adjoint representation") of the algebra from

[29] I borrow here from what is called "Campbell-Baker-Hausdorff theory." See CLASSICAL MECHANICS (1980), p. 285 and references cited there. But we need not venture into that intricate topic to understand the simple point at issue.

which they sprang.[30] Look, for example, to the group $SU(2)$: we are led from the structure constants implicit in (76.3) to the matrices

$$\mathbb{C}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -2i \\ 0 & +2i & 0 \end{pmatrix},\ \mathbb{C}_2 = \begin{pmatrix} 0 & 0 & +2i \\ 0 & 0 & 0 \\ -2i & 0 & 0 \end{pmatrix},\ \mathbb{C}_3 = \begin{pmatrix} 0 & -2i & 0 \\ +2i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Calculation confirms that these matrices do in fact satisfy the commutation relations (76.3), even though the set $\{\mathbb{C}_1, \mathbb{C}_2, \mathbb{C}_3\}$ is not multiplicatively closed and therefore cannot possibly satisfy (76.1/2).

The $\mathcal{N} \times \mathcal{N}$ matrix

$$\Bbbk \equiv \|k_{pq}\| \quad \text{with} \quad k_{pq} \equiv \operatorname{tr} \mathbb{C}_p \mathbb{C}_q \tag{82}$$

is transparently real and symmetric. Proceeding in the assumption that $\det \Bbbk \neq 0$, I write $\Bbbk^{-1} \equiv \|k^{pq}\|$ and will assign to $\Bbbk$ the role of "gauge metric;" i.e., I will use k^{pq} and k_{pq} to raise and lower indices. In the case $SU(2)$ we compute

$$\Bbbk = \begin{pmatrix} \operatorname{tr}\mathbb{C}_1\mathbb{C}_1 & \operatorname{tr}\mathbb{C}_1\mathbb{C}_2 & \operatorname{tr}\mathbb{C}_1\mathbb{C}_3 \\ \operatorname{tr}\mathbb{C}_2\mathbb{C}_1 & \operatorname{tr}\mathbb{C}_2\mathbb{C}_2 & \operatorname{tr}\mathbb{C}_2\mathbb{C}_3 \\ \operatorname{tr}\mathbb{C}_3\mathbb{C}_1 & \operatorname{tr}\mathbb{C}_3\mathbb{C}_2 & \operatorname{tr}\mathbb{C}_3\mathbb{C}_3 \end{pmatrix} = \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

The results just obtained can be interpreted to state that the traceless hermitian 3×3 matrices $\mathbb{E}_p \equiv \frac{1}{\sqrt{8}}\mathbb{C}_p$ are *tracewise orthonormal*, and are *special to the Pauli basis*; with respect to that basis one has $k_{pq} = 8\delta_{pq}$.

We will have need of a result which can be stated

$$C_{psq} \text{ is totally antisymmetric} \tag{83.1}$$

and which I digress now to establish. We have

$$\begin{aligned} C_{psq} = k_{sr} \cdot C_p{}^r{}_q = C_s{}^u{}_v \underbrace{C_r{}^v{}_u \cdot C_p{}^r{}_q} & \\ = C_p{}^v{}_w C_q{}^w{}_u - C_q{}^v{}_w C_p{}^w{}_u & \quad \text{by (80.2)} \\ = C_s{}^u{}_v C_p{}^v{}_w C_q{}^w{}_u + C_v{}^u{}_s C_u{}^w{}_p C_w{}^v{}_q & \quad \text{by (80.1)} \\ = \left\{ \begin{array}{l} \text{sum or terms each of which is invariant} \\ \text{under cyclic permutation on } \{spq\} \end{array} \right. & \end{aligned}$$

from which we conclude that $C_{psq} = C_{sqp} = C_{qps}$. But $C_{psq} = -C_{qsp}$. This establishes (83.1), from which it follows as a useful corollary that

$$C_{psr} = -C_{prs} \quad \text{which is to say:} \quad (\Bbbk\, \mathbb{C}_p)^{\intercal} = -(\Bbbk\, \mathbb{C}_p) \tag{83.2}$$

[30] For discussion which digs deeper into the relevant group theory, see §5.6 in H. Bacry, *Lectures on Group Theory and Particle Theory* (1977).

Finally a word about notation: one designs notation so as to be in position to say simple things simply, to highlight essentials while not masking critical distinctions. In classical non-Abelian gauge theory only simple things are going on (some linear algebra, some elementary calculus), but they are going on in potentially confusing constellation. To write gauge field theory in explicit detail would bring into play such blizzard of indices (of diverse ranges and meanings) as to make it very difficult to gain a sense of *what is going on*. But to surpress such detail—to adopt the scrubbed notation standard to publication in the field—is to risk losing a vivid sense of what the marks on the page specifically *mean*. My purpose here is to point out that classical mathematics does supply a tool which in this instance permits one to strike a happy medium; the tool has a name, but it is a name seldom encountered in the gauge field theoretic literature.

Suppose, by way of introduction, that we have interest in a pair of 3-vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, which we propose first to subject independently to linear transformations, and then to rotationally intermix; we might write[31]

$$\left.\begin{aligned} \boldsymbol{x} &\longrightarrow \mathbb{A}\,\boldsymbol{x} \\ \boldsymbol{y} &\longrightarrow \mathbb{B}\,\boldsymbol{y} \end{aligned}\right\} \longrightarrow \left\{\begin{aligned} &\cos\theta\cdot\mathbb{A}\,\boldsymbol{x} - \sin\theta\cdot\mathbb{B}\,\boldsymbol{y} \\ &\sin\theta\cdot\mathbb{A}\,\boldsymbol{x} + \cos\theta\cdot\mathbb{B}\,\boldsymbol{y} \end{aligned}\right.$$

But if we "stack" the 3-vectors (forming a 6-vector) we acquire this alternative means of displaying the same information:

$$\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix} \longrightarrow \begin{pmatrix}\mathbb{A} & \mathbb{O}\\ \mathbb{O} & \mathbb{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix} \longrightarrow \begin{pmatrix}\cos\theta\cdot\mathbb{I} & -\sin\theta\cdot\mathbb{I}\\ \sin\theta\cdot\mathbb{I} & \cos\theta\cdot\mathbb{I}\end{pmatrix}\begin{pmatrix}\mathbb{A} & \mathbb{O}\\ \mathbb{O} & \mathbb{B}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix}$$

The "Kronecker product" (sometimes called the "direct product") of
- an $m \times n$ matrix $\mathbb{A}$ onto
- a $p \times q$ matrix $\mathbb{B}$

is the $mp \times nq$ matrix defined[32]

$$\mathbb{A} \otimes \mathbb{B} \equiv \|a_{ij}\,\mathbb{B}\| \tag{84}$$

In that notation, the "matrices with matrix-valued elements" encountered in my example can be described

$$\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix} \otimes \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

[31] For the purposes of this discussion I revert to my former practice of writing vectors in boldface, matrices in blackboard doubleface.

[32] The alternative definition $\mathbb{A} \otimes \mathbb{B} \equiv \|\mathbb{A}\,b_{ij}\|$ gives rise to a "mirror image" of the standard theory. Good discussions can be found in E. P. Wigner, *Group Theory and its Application to the Quantum Theory of Atomic Spectra* (1959), Chapter 2; P. Lancaster, *Theory of Matrices* (1969), §8.2; Richard Bellman, *Introduction to Matrix Analysis* (2nd edition 1970), Chapter 12, §§5–13.

and

$$\begin{pmatrix}1&0\\0&0\end{pmatrix}\otimes\begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}\quad+\quad\begin{pmatrix}0&0\\0&1\end{pmatrix}\otimes\begin{pmatrix}b_{11}&b_{12}&b_{13}\\b_{21}&b_{22}&b_{23}\\b_{31}&b_{32}&b_{33}\end{pmatrix}$$

Manipulation of expressions involving Kronecker products is accomplished by appeal to general statements such as the following:

$$k(\mathbb{A}\otimes\mathbb{B})=(k\mathbb{A})\otimes\mathbb{B}=\mathbb{A}\otimes(k\mathbb{B}) \tag{85.1}$$

$$\left.\begin{aligned}(\mathbb{A}+\mathbb{B})\otimes\mathbb{C}&=\mathbb{A}\otimes\mathbb{C}+\mathbb{B}\otimes\mathbb{C}\\ \mathbb{A}\otimes(\mathbb{B}+\mathbb{C})&=\mathbb{A}\otimes\mathbb{B}+\mathbb{A}\otimes\mathbb{C}\end{aligned}\right\} \tag{85.2}$$

$$\mathbb{A}\otimes(\mathbb{B}\otimes\mathbb{C})=(\mathbb{A}\otimes\mathbb{B})\otimes\mathbb{C}\equiv\mathbb{A}\otimes\mathbb{B}\otimes\mathbb{C} \tag{85.3}$$

$$(\mathbb{A}\otimes\mathbb{B})^{\intercal}=\mathbb{A}^{\intercal}\otimes\mathbb{B}^{\intercal} \tag{85.4}$$

$$\mathrm{tr}(\mathbb{A}\otimes\mathbb{B})=\mathrm{tr}\mathbb{A}\cdot\mathrm{tr}\mathbb{B} \tag{85.5}$$

—all of which are valid except when meaningless.[33] Less obviously (but often very usefully)

$$(\mathbb{A}\otimes\mathbb{B})(\mathbb{C}\otimes\mathbb{D})=\mathbb{A}\mathbb{C}\otimes\mathbb{B}\mathbb{D}\quad\text{if}\quad\begin{cases}\mathbb{A}\text{ and }\mathbb{C}\text{ are }m\times m\\ \mathbb{B}\text{ and }\mathbb{D}\text{ are }n\times n\end{cases} \tag{85.6}$$

from which one can extract[34]

$$\mathbb{A}\otimes\mathbb{B}=(\mathbb{A}\otimes\mathbb{I}_n)(\mathbb{I}_m\otimes\mathbb{B}) \tag{85.7}$$

$$\det(\mathbb{A}\otimes\mathbb{B})=(\det\mathbb{A})^n(\det\mathbb{B})^m \tag{85.8}$$

$$(\mathbb{A}\otimes\mathbb{B})^{-1}=\mathbb{A}^{-1}\otimes\mathbb{B}^{-1} \tag{85.9}$$

Here I have used $\mathbb{I}_m$ to designate the $m\times m$ identity matrix, and below use $\boldsymbol{I}_m$ for that same purpose (though when the dimension is obvious from the context I allow myself to omit the subscript).

Dirac theory with local SU(2) gauge invariance. I look now to the historic first instance of a "non-Abelian gauge field theory"—the theory put forward by Yang & Mills (though presented here as an exercise in *classical* field theory). We select Dirac theory as our starting point because it is, in many respects, "simplest possible." And we select $SU(2)$ as our non-Abelian gauge group for that same reason (Yang & Mills had their own—more pressingly physical—reasons for both selections).

[33] Recall that one cannot add matrices unless they are co-dimensional, and does not speak of the trace of a matrix unless it is square.

[34] See Lancaster[32] for the detailed arguments.

Our first assignment STEP ZERO is to construct a variant of the Dirac theory which exhibits *global* $SU(2)$ invariance. To that end we "stack" two copies of the familiar Dirac theory: we assemble an 8-component complex field

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} \quad \text{with} \quad \psi^a = \begin{pmatrix} \psi^a_1 \\ \psi^a_2 \\ \psi^a_3 \\ \psi^a_4 \end{pmatrix} \quad : \quad a = 1, 2 \tag{86}$$

(for lack of standard terminology I will call the superscripts "gauge indices" and the subscripts "Dirac indices") and require that it satisfy the field equations

$$\big(\boldsymbol{\Gamma}^\mu \partial_\mu + i\boldsymbol{\varkappa}\big)\psi = 0 \quad \text{and adjoint} \tag{87}$$

where

$$\boldsymbol{\Gamma}^\mu \equiv \boldsymbol{I}_2 \otimes \boldsymbol{\gamma}^\mu \quad \text{and} \quad \boldsymbol{\varkappa} \equiv \begin{pmatrix} \varkappa_1 & 0 \\ 0 & \varkappa_2 \end{pmatrix} \otimes \boldsymbol{I}_4 \tag{88}$$

The field equations arise from

$$\mathcal{L}_0(\psi, \tilde{\psi}, \partial\psi, \partial\tilde{\psi}) = \hbar c \Big[\tfrac{1}{2} i \big\{ \tilde{\psi} \boldsymbol{\Gamma}^\alpha \psi_{,\alpha} - \tilde{\psi}_{,\alpha} \boldsymbol{\Gamma}^\alpha \psi \big\} - \tilde{\psi} \boldsymbol{\varkappa} \psi \Big] \tag{89}$$

which—because the matrices $\boldsymbol{\Gamma}^\mu$ and $\boldsymbol{\varkappa}$ share the block structure

$$\begin{pmatrix} \bullet & \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 \\ \bullet & \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 \\ \bullet & \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 \\ \bullet & \bullet & \bullet & \bullet & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet \\ 0 & 0 & 0 & 0 & \bullet & \bullet & \bullet & \bullet \end{pmatrix}$$

—is just the sum of the Dirac Lagrangians which separately regulate the (presently) uncoupled motion of ψ^1 and ψ^2.

The 8×8 matrices $\boldsymbol{S} \equiv \boldsymbol{s} \otimes \boldsymbol{I}_4$ mimic the multiplicative properties of the 2×2 unimodular unitary matrices $\boldsymbol{s}$

$$\begin{aligned} \boldsymbol{S}_2 \boldsymbol{S}_1 &= (\boldsymbol{s}_2 \otimes \boldsymbol{I}_4)(\boldsymbol{s}_1 \otimes \boldsymbol{I}_4) \\ &= \boldsymbol{s}_2 \boldsymbol{s}_1 \otimes \boldsymbol{I}_4 \quad \text{by (85.6)} \end{aligned}$$

and are readily seen to be themselves unimodular and unitary.[35] Corresponding to the "Pauli decomposition"

$$\boldsymbol{s} = s^\mu \boldsymbol{\sigma}_\mu$$

we have

$$\boldsymbol{S} = s^\mu \boldsymbol{\Sigma}_\mu \quad \text{with} \quad \boldsymbol{\Sigma}_\mu \equiv \boldsymbol{\sigma}_\mu \otimes \boldsymbol{I}_4 \tag{90}$$

[35] Use (85.4), (85.8) and (85.9).

Specifically

$$\boldsymbol{\Sigma}_0 \equiv \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I} \end{pmatrix},\ \boldsymbol{\Sigma}_1 \equiv \begin{pmatrix} \boldsymbol{0} & \boldsymbol{I} \\ \boldsymbol{I} & \boldsymbol{0} \end{pmatrix},\ \boldsymbol{\Sigma}_2 \equiv \begin{pmatrix} \boldsymbol{0} & -\boldsymbol{i} \\ \boldsymbol{i} & \boldsymbol{0} \end{pmatrix},\ \boldsymbol{\Sigma}_3 \equiv \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I} \end{pmatrix}$$

where the sub-matrices are 4×4. Note particularly that $\boldsymbol{\Sigma}_1$ and $\boldsymbol{\Sigma}_2$ have the reverse of the block structure pictured above.

From the reversed block structure just mentioned it follows that the Lagrangian (89) will be (manifestly) invariant under $\psi \longrightarrow \boldsymbol{S}\psi$ if and only if $\varkappa_1 = \varkappa_2$, which we will assume.[36] $SU(2)$ is a 3-parameter group, so from the *global* $SU(2)$-invariance of $\mathcal{L}_0$—just established— follow three conservation laws. Infinitesimally, we have

$$\begin{aligned} \psi &\longrightarrow \psi + \delta\psi \\ \delta\psi &= i\boldsymbol{\Sigma}_1\psi \cdot \delta\omega_1 + i\boldsymbol{\Sigma}_2\psi \cdot \delta\omega_2 + i\boldsymbol{\Sigma}_3\psi \cdot \delta\omega_3 \end{aligned}$$

and so are led by Noether to the statements (compare (69))

$$\partial_\mu Q^\mu_r = 0 \quad \text{with} \quad Q^\mu_r \equiv \tfrac{1}{2}c\big\{\tilde{\psi}\boldsymbol{\Gamma}^\mu\boldsymbol{\Sigma}_r\psi + \text{conjugate}\big\} \quad : \quad r = 1, 2, 3$$

Letting the $\boldsymbol{\Sigma}$-matrices, as described above, act upon $\begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix}$ we obtain

$$\boldsymbol{\Sigma}_1\begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} = \begin{pmatrix} \psi^2 \\ \psi^1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_2\begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} = -i\begin{pmatrix} \psi^2 \\ -\psi^1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_3\begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} = \begin{pmatrix} \psi^1 \\ -\psi^2 \end{pmatrix}$$

which yield these more explicit descriptions of the conserved currents Q^μ_r:

$$Q^\mu_1 = \quad \tfrac{1}{2}c\big(\tilde{\psi}^1\boldsymbol{\gamma}^\mu\psi^2 + \tilde{\psi}^2\boldsymbol{\gamma}^\mu\psi^1\big) \tag{91.1}$$

$$Q^\mu_2 = -i\tfrac{1}{2}c\big(\tilde{\psi}^1\boldsymbol{\gamma}^\mu\psi^2 - \tilde{\psi}^2\boldsymbol{\gamma}^\mu\psi^1\big) \tag{91.2}$$

$$Q^\mu_3 = \quad \tfrac{1}{2}c\big(\tilde{\psi}^1\boldsymbol{\gamma}^\mu\psi^1 - \tilde{\psi}^2\boldsymbol{\gamma}^\mu\psi^2\big) \tag{91.3}$$

The twinned Dirac Lagrangian $\mathcal{L}_0$ is also (manifestly) U(1)-invariant, which leads to conservation of

$$Q^\mu_0 = \quad \tfrac{1}{2}c\big(\tilde{\psi}^1\boldsymbol{\gamma}^\mu\psi^1 + \tilde{\psi}^2\boldsymbol{\gamma}^\mu\psi^2\big) \tag{91.0}$$

which is the anticipated twinned instance of (69.2).[37]

---

[36] At this point Yang & Mills, following in the footsteps of the inventors of the isotopic spin concept (Heisenberg, Wigner and others, in the late 1930's), were content to draw upon the physical circumstance that the proton and neutron masses are *nearly* the same:

$$m_p = 938.280 \text{ MeV}/c^2 \quad \text{while} \quad m_n = 939.573 \text{ MeV}/c^2$$

[37] The quartet of conservation laws (90) are structurally reminiscent of a quartet encountered in connection with the classical mechanics of an isotropic 2-dimensional oscillator. Nor is that formal connection surprising: here $SU(2)$ is an explicitly imposed symmetry; there it is a "hidden symmetry." See the discussion surrounding equation (162) in my "Ellipsometry" (1999).

Our objective is to achieve *local* $SU(2)$ invariance, and we confront at the outset the familiar problem that

$$\psi \longrightarrow \psi' = \boldsymbol{S}(x)\psi \quad \text{induces} \quad \psi_{,\mu} \longrightarrow \psi'_{,\mu} = \boldsymbol{S}(x)\psi_{,\mu} + \boldsymbol{S}_{,\mu}(x)\psi$$

It is to escape the force of the elementary circumstance that ψ and $\psi_{,\mu}$ transform by different rules that STEP ONE we make what we have learned to call the "minimal coupling substitution"[38]

$$\begin{aligned} &\boldsymbol{\partial}_\mu \\ &\downarrow \\ &\boldsymbol{\mathcal{D}}_\mu = \boldsymbol{\partial}_\mu - ig\boldsymbol{A}_\mu \quad \text{with} \quad g \equiv e/\hbar c \end{aligned} \tag{92}$$

and STEP TWO concoct $\boldsymbol{A}_\mu \rightarrow \boldsymbol{A}'_\mu$ so as to achieve $(\boldsymbol{\mathcal{D}}_\mu\psi)' = \boldsymbol{S}(x)(\boldsymbol{\mathcal{D}}_\mu\psi)$. From

$$\Big[(\boldsymbol{\partial}_\mu - ig\boldsymbol{A}'_\mu)\boldsymbol{S} = \boldsymbol{S}\boldsymbol{\partial}_\mu + \boldsymbol{S}_{,\mu} - ig\boldsymbol{A}'_\mu\boldsymbol{S}\Big] = \boldsymbol{S}(\boldsymbol{\partial}_\mu - ig\boldsymbol{A}_\mu)$$

we are led thus to this enlarged interpretation

$$\left.\begin{aligned} \psi \;&\longrightarrow \psi' = \boldsymbol{S}\,\psi \\ \boldsymbol{A}_\mu &\longrightarrow \boldsymbol{A}'_\mu = \boldsymbol{S}\boldsymbol{A}_\mu\boldsymbol{S}^{-1} + i\,\tfrac{1}{g}\,\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1} \end{aligned}\right\} \tag{93}$$

of what we shall understand the phrase "local $SU(2)$ gauge transformation" to *mean*.

Equation (93) describes the non-Abelian counterpart to (73), and the points of similarity/difference stand out even more clearly when we write

$$\boldsymbol{S} = e^{ig\boldsymbol{\Omega}} \quad : \quad \boldsymbol{\Omega} \text{ traceless hermitian}$$

Whereas the $A_\mu \longrightarrow A'_\mu = A_\mu + \partial_\mu\Omega$ encountered in the Abelian case $U(1)$ had the form of a

derivative-dependent shift

its non-Abelian counterpart in (93) has the form

similarity transformation + derivative-dependent shift

It is non-commutivity ($[\boldsymbol{S}, \boldsymbol{A}_\mu] \neq \boldsymbol{0}$) which prevents the $\boldsymbol{S}$ from slipping past the $\boldsymbol{A}_\mu$ and cancelling the $\boldsymbol{S}^{-1}$; i.e., which accounts for the survival of the similarity transformation as a characteristic component of non-Abelian gauge.

Differentiation of the unitarity condition $\boldsymbol{S}\boldsymbol{S}^\dagger = \boldsymbol{I}$ supplies the information that $i\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1}$ is invariably hermitian. And, of course, $\boldsymbol{S}\boldsymbol{A}_\mu\boldsymbol{S}^{-1}$ is (traceless) hermitian if $\boldsymbol{A}_\mu$ is. We therefore *assume the gauge matrices* $\boldsymbol{A}_\mu$ *to be hermitian*, and observe it to be an implication of (93) that

$$\boldsymbol{A}_\mu \longrightarrow \boldsymbol{A}'_\mu \text{ preserves hermiticity}$$

The matrices $\boldsymbol{S}$ are, however, not just *any* old unimodular unitary 8×8 matrices; they possess the specialized structure $\boldsymbol{S} \equiv \boldsymbol{s} \otimes \boldsymbol{I}_4$, reflecting the fact that our interest at the moment lies not in $SU(8)$ but in $SU(2)$. We impute that

[38] It is fussy of me to write $\boldsymbol{\partial}_\mu$ (meaning $\boldsymbol{I}\partial_\mu$) in place more simply of ∂_μ, but it offends my eye to "add a scalar to a matrix."

structure also to the hermitian matrices $\boldsymbol{A}_\mu$, writing

$$\boldsymbol{A}_\mu \equiv \boldsymbol{a}_\mu \otimes \boldsymbol{I}_4 \quad \text{with } \boldsymbol{a}_\mu \text{ } 2\times 2 \text{ hermitian}$$

From the unimodularity of $\boldsymbol{S} = e^{ig\boldsymbol{\Omega}}$ we know that (as previously remarked) $\boldsymbol{\Omega}$ is necessarily traceless, and can therefore be developed

$$\boldsymbol{\Omega} = \qquad \Omega^1\boldsymbol{\Sigma}_1 + \Omega^2\boldsymbol{\Sigma}_2 + \Omega^3\boldsymbol{\Sigma}_3$$

But we might expect to have to write

$$\boldsymbol{A}_\mu = A^0_\mu\boldsymbol{\Sigma}_0 + A^1_\mu\boldsymbol{\Sigma}_1 + A^2_\mu\boldsymbol{\Sigma}_2 + A^3_\mu\boldsymbol{\Sigma}_3$$

I will argue that the ^{0} term can be abandoned; i.e., that one can without loss of generality *assume the gauge matrices $\boldsymbol{A}_\mu$ to be traceless*. The argument proceeds in two steps, of which the first, unfortunately, is a bit intricate: we demonstrate that the traceless assumption—if made—is transformationally stable. Certainly it is the case[39] that if $\boldsymbol{A}_\mu$ is traceless then so is $\boldsymbol{S}\boldsymbol{A}_\mu\boldsymbol{S}^{-1}$. But how to show that $\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1}$ is traceless? I quote two general identities[40]

$$\begin{aligned} e^{ig\boldsymbol{\Omega}}\boldsymbol{A}\,e^{-ig\boldsymbol{\Omega}} &= \left\{e^{ig\boldsymbol{\Omega}}, \boldsymbol{A}\right\} \\ &\equiv \boldsymbol{A} + ig[\boldsymbol{\Omega}, \boldsymbol{A}] + \tfrac{1}{2!}(ig)^2[\boldsymbol{\Omega}, [\boldsymbol{\Omega}, \boldsymbol{A}]] + \cdots \end{aligned} \tag{94.1}$$

$$\begin{aligned} \frac{\partial e^{ig\boldsymbol{\Omega}}}{\partial x}\cdot e^{-ig\boldsymbol{\Omega}} &= \left\{\frac{e^{ig\boldsymbol{\Omega}} - \boldsymbol{I}}{\boldsymbol{\Omega}}, \tfrac{\partial}{\partial x}\boldsymbol{\Omega}\right\} \\ &= ig\tfrac{\partial}{\partial x}\boldsymbol{\Omega} + \tfrac{1}{2!}(ig)^2[\boldsymbol{\Omega}, \tfrac{\partial}{\partial x}\boldsymbol{\Omega}] + \tfrac{1}{3!}(ig)^3[\boldsymbol{\Omega}, [\boldsymbol{\Omega}, \tfrac{\partial}{\partial x}\boldsymbol{\Omega}]] + \cdots \end{aligned} \tag{94.2}$$

but have immediate need only of the second. Clearly $\frac{\partial}{\partial x}\boldsymbol{\Omega}$ is traceless if $\boldsymbol{\Omega}$ is. But it was remarked already in connection with (79) that

$$[\text{hermitian}, \text{hermitian}] = i\,(\text{traceless hermitian})$$

so each of the nested commutators presented on the right side of (94.2) are in fact traceless. This establishes the point at issue: if $\boldsymbol{A}_\mu$ is traceless then so, according to (93), is $\boldsymbol{A}'_\mu$. Completion of the argument must await one further development:

Given our initial interest in the Dirac Lagrangian (which at (89) we duplicated, in order to get this show on the road), we STEP THREE look to the modified system

$$\mathcal{L}_1(\psi, \tilde{\psi}, \partial\psi, \partial\tilde{\psi}, A^1, A^2, A^3) = \mathcal{L}_0(\psi, \tilde{\psi}, \mathcal{D}\psi, \tilde{\mathcal{D}}\tilde{\psi})$$

---

[39] Use $\text{tr}(\mathbb{A}\mathbb{B}) = \text{tr}(\mathbb{B}\mathbb{A})$.

[40] See, for example, §4 in R. M. Wilcox, "Exponential operators & parameter differentiation in quantum physics," J. Math. Phys. **8**, 962 (1967). The identities in question are attributed by W. Magnus (in "On the exponential solution of differential equations for a linear operator," Comm. Pure & Appl. Math. **7**, 649 (1954)) to F. Hausdorff (1906). In (93.1) one is free to install subscripts on all the $\boldsymbol{A}$'s, while in (93.2) one can assign any meaning to the parameter x.

In more explicit detail we have

$$\begin{aligned}\mathcal{L}_1 &= \hbar c\Big[\tfrac{1}{2}i\Big\{\tilde{\psi}\boldsymbol{\gamma}^\alpha(\psi_{,\alpha} - i\tfrac{e}{\hbar c}\boldsymbol{A}_\alpha\psi) - (\tilde{\psi}_{,\alpha} + i\tfrac{e}{\hbar c}\tilde{\psi}\boldsymbol{A}_\alpha)\boldsymbol{\gamma}^\alpha\psi\Big\} - \varkappa\tilde{\psi}\psi\Big] \\ &= \hbar c\Big[\tfrac{1}{2}i\Big\{\tilde{\psi}\boldsymbol{\gamma}^\alpha\big[\psi_{,\alpha} - i\tfrac{e}{\hbar c}(A^0_\alpha\boldsymbol{\Sigma}_0 + A^p_\alpha\boldsymbol{\Sigma}_p)\psi\big] \\ &\qquad\qquad - \big[\tilde{\psi}_{,\alpha} + i\tfrac{e}{\hbar c}\tilde{\psi}(A^0_\alpha\boldsymbol{\Sigma}_0 + A^p_\alpha\boldsymbol{\Sigma}_p)\big]\boldsymbol{\gamma}^\alpha\psi\Big\} - \varkappa\tilde{\psi}\psi\Big]\end{aligned} \tag{95}$$

We observe that an $A^0_\mu\boldsymbol{\Sigma}_0$ term, if assumed to be present in the composition of $\boldsymbol{A}_\mu$, would (since $\boldsymbol{\Sigma}_0 = \boldsymbol{I}_8$ commutes with everything) simply replicate the adjustment we would make—and *did* make at (70)—if we were trying to achieve $U(1)$ gauge invariance. We will agree to abandon the hypothetical ^{0} term on grounds that it is passive with respect to $SU(2)$, and that its discovered predisposition is to talk about something (electrodynamics) other than the subject that presently interests us.

So we have

$$\boldsymbol{A}_\mu(x) = A^1_\mu(x)\boldsymbol{\Sigma}_1 + A^2_\mu(x)\boldsymbol{\Sigma}_2 + A^3_\mu(x)\boldsymbol{\Sigma}_3 \tag{96}$$

and at this point make the acquaintance of the *three "gauge fields"—one for each generator—called into being by the imposition of local $SU(3)$ invariance.* It is important to notice that the gauge fields $A^p_\mu(x)$ are necessarily *real-valued vector* fields, and that they arise as "coordinates" of the more fundamental objects $\boldsymbol{A}_\mu(x)$: select a different basis (in the space of traceless hermitian 2×2 matrices) and be led from the same $\boldsymbol{A}_\mu(x)$ to a different trio of gauge fields. The matrix-valued gauge field $\boldsymbol{A}_\mu(x)$ cannot be accorded "physical immediacy" because susceptible to gauge, and its coordinates have an even more tenuous claim to reality.

We have now on-stage a total of twelve real-number-valued fields

$$\begin{matrix} A^1_0(x) & A^1_1(x) & A^1_2(x) & A^1_3(x) \\ A^2_0(x) & A^2_1(x) & A^2_2(x) & A^2_3(x) \\ A^3_0(x) & A^3_1(x) & A^3_2(x) & A^3_3(x) \end{matrix} \tag{97}$$

Lorentz transformations linearly recombine the columns; gauge transformations linearly recombine the rows. It is instructive to inquire into details of the latter process. To that end: observe in connection with (94.1), which presents a "Lie series" (i.e., a series of nested commutators) on its right hand side, that

$$\begin{aligned}[\boldsymbol{\Omega},\boldsymbol{A}\,] &= \Omega^p A^q[\,\boldsymbol{\Sigma}_p,\boldsymbol{\Sigma}_q\,] \quad \text{with } \sum_{p,q=1}^{3} \text{ here as henceforth understood} \\ &= \Omega^p A^q\, i{c_p}^r{}_q\,\boldsymbol{\Sigma}_r\end{aligned} \tag{98.1}$$

and that from the $SU(2)$ commutation relations (76.3) one has, in the Pauli basis,

$${c_p}^r{}_q = 2\,\mathrm{sgn}\tbinom{1\,2\,3}{p\,q\,r} \tag{98.2}$$

which, by the way, conforms nicely to (83.1).[41] Returning with (98.2) to (98.1), we obtain

$$\begin{aligned}
[\boldsymbol{\Omega}, \boldsymbol{A}\,] &= 2i\,(\vec{\Omega} \times \vec{A}\,)^r\,\boldsymbol{\Sigma}_r \\
&= \left[\underbrace{\begin{pmatrix} 0 & -2i\Omega^3 & 2i\Omega^2 \\ 2i\Omega^3 & 0 & -2i\Omega^1 \\ -2i\Omega^2 & 2i\Omega^1 & 0 \end{pmatrix}}_{\mathbb{Z}} \begin{pmatrix} A^1 \\ A^2 \\ A^3 \end{pmatrix} \right] \cdot \begin{pmatrix} \boldsymbol{\Sigma}_1 \\ \boldsymbol{\Sigma}_2 \\ \boldsymbol{\Sigma}_3 \end{pmatrix} \\
&= (\mathbb{Z}\vec{A}\,)^r\,\boldsymbol{\Sigma}_r \\
[\boldsymbol{\Omega}, [\boldsymbol{\Omega}, \boldsymbol{A}\,]] &= (\mathbb{Z}^2\vec{A}\,)^r\,\boldsymbol{\Sigma}_r \\
&\;\;\vdots
\end{aligned}$$

Returning with this information to (94) we find that the second half of (93) can be rendered

$$\vec{A}_\mu \longrightarrow \vec{A}'_\mu = e^{ig\mathbb{Z}}\vec{A}_\mu - \frac{e^{ig\mathbb{Z}} - \mathbb{I}}{ig\mathbb{Z}}\,\vec{\Omega}_{,\mu} \tag{99.1}$$

Looking now with one eye to the definition of $\mathbb{Z}$ and with the other to the equations which (just prior to (82)) served to define $\{\mathbb{C}_1, \mathbb{C}_2, \mathbb{C}_3\}$, we see that

$$\mathbb{Z} = \Omega^1\mathbb{C}_1 + \Omega^2\mathbb{C}_2 + \Omega^3\mathbb{C}_3 \tag{100}$$

and notice, moreover, that

$$ig\mathbb{Z} = g\begin{pmatrix} 0 & 2\Omega^3 & -2\Omega^2 \\ -2\Omega^3 & 0 & 2\Omega^1 \\ 2\Omega^2 & -2\Omega^1 & 0 \end{pmatrix} \quad \text{is real antisymmetric}$$

so

$$\mathbb{R}(x) \equiv e^{ig\mathbb{Z}(x)} \quad \text{is a rotation matrix, an element of } O(3)$$

In this notation (99.1) can be rendered

$$\vec{A}_\mu \longrightarrow \vec{A}'_\mu = \mathbb{R}\vec{A}_\mu - \frac{\mathbb{R} - \mathbb{I}}{\log\mathbb{R}}\,\vec{\Omega}_{,\mu} \tag{99.2}$$

Equations (99) say the same thing, the latter being a "coordinatized" version of the former. Note the natural occurance of the adjoint representation at (98).

We have now STEP FOUR to prepare to lauch the gauge matrix $\boldsymbol{A}_\mu$ (equivalently: the gauge fields $A^p_\mu : p = 1, 2, 3$) into dynamical motion. To that

[41] One should resist the temptation to write $c_p{}^r{}_q = 2\epsilon_{pqr}$, for although the equation is *numerically* correct *in the Pauli basis* it is transformationally screwy: it presents ^{r} on the left but $_r$ on the right, and would come unstuck if one were to abandon the Pauli basis in favor of some arbitrary alternative.

end, taking (15) as our model, we examine the gauge transformation properties of

$$\boldsymbol{f}_{\mu\nu} = \partial_\mu \boldsymbol{A}_\nu - \partial_\nu \boldsymbol{A}_\mu$$

From (93) we obtain

$$\begin{aligned}\partial_\mu \boldsymbol{A}_\nu \longrightarrow \partial_\mu \boldsymbol{A}'_\nu = {} & \boldsymbol{S}_{,\mu}\boldsymbol{A}_\nu\boldsymbol{S}^{-1} + \boldsymbol{S}(\partial_\mu \boldsymbol{A}_\nu)\boldsymbol{S}^{-1} + \boldsymbol{S}\boldsymbol{A}_\nu(\boldsymbol{S}^{-1})_{,\mu} \\ & + i\,\tfrac{1}{g}\,\boldsymbol{S}_{,\nu\mu}\boldsymbol{S}^{-1} + i\,\tfrac{1}{g}\,\boldsymbol{S}_{,\nu}(\boldsymbol{S}^{-1})_{,\mu}\end{aligned}$$

But $\partial_\mu(\boldsymbol{S}\boldsymbol{S}^{-1}) = \boldsymbol{0}$ supplies $(\boldsymbol{S}^{-1})_{,\mu} = -\boldsymbol{S}^{-1}\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1}$ so

$$\begin{aligned} = {} & \boldsymbol{S}(\partial_\mu \boldsymbol{A}_\nu)\boldsymbol{S}^{-1} \\ & + \boldsymbol{S}_{,\mu}\boldsymbol{A}_\nu\boldsymbol{S}^{-1} - \boldsymbol{S}\boldsymbol{A}_\nu\boldsymbol{S}^{-1}\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1} + i\tfrac{1}{g}\,\boldsymbol{S}_{,\mu\nu}\boldsymbol{S}^{-1} - i\tfrac{1}{g}\,\boldsymbol{S}_{,\nu}\boldsymbol{S}^{-1}\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1}\end{aligned}$$

gives

$$\boldsymbol{f}_{\mu\nu} \longrightarrow \boldsymbol{f}'_{\mu\nu} = \boldsymbol{S}\boldsymbol{f}_{\mu\nu}\boldsymbol{S}^{-1} + \big\{\text{unwelcome term}\big\}$$

with

$$\begin{aligned}\big\{\text{unwelcome term}\big\} = {} & (\boldsymbol{S}_{,\mu}\boldsymbol{A}_\nu\boldsymbol{S}^{-1} - \boldsymbol{S}_{,\nu}\boldsymbol{A}_\mu\boldsymbol{S}^{-1}) \\ & + (\boldsymbol{S}\boldsymbol{A}_\mu\boldsymbol{S}^{-1}\boldsymbol{S}_{,\nu}\boldsymbol{S}^{-1} - \boldsymbol{S}\boldsymbol{A}_\nu\boldsymbol{S}^{-1}\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1}) \\ & + i\tfrac{1}{g}(\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1}\boldsymbol{S}_{,\nu}\boldsymbol{S}^{-1} - \boldsymbol{S}_{,\nu}\boldsymbol{S}^{-1}\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1})\end{aligned}$$

The non-commutivity responsible for the existence of the "unwelcome term" is responsible also for the existence of a second $\mu\nu$-antisymmetric construct—namely the commutator $[\boldsymbol{A}_\mu, \boldsymbol{A}_\nu]$, which is found by straightforward calculation to transform

$$[\boldsymbol{A}_\mu, \boldsymbol{A}_\nu] \longrightarrow [\boldsymbol{A}'_\mu, \boldsymbol{A}'_\nu] = \boldsymbol{S}[\boldsymbol{A}_\mu, \boldsymbol{A}_\nu]\boldsymbol{S}^{-1} - i\tfrac{1}{g}\big\{\text{same unwelcome term}\big\}$$

The pretty implication is that

$$\boldsymbol{F}_{\mu\nu} \equiv (\partial_\mu \boldsymbol{A}_\nu - \partial_\nu \boldsymbol{A}_\mu) - ig(\boldsymbol{A}_\mu\boldsymbol{A}_\nu - \boldsymbol{A}_\nu\boldsymbol{A}_\mu) \tag{101}$$

gauge-transforms by simple similarity transformation (since the "unwelcome terms" cancel):

$$\boldsymbol{F}_{\mu\nu} \longrightarrow \boldsymbol{F}'_{\mu\nu} = \boldsymbol{S}\boldsymbol{F}_{\mu\nu}\boldsymbol{S}^{-1} \tag{102.1}$$

Which (to say the same thing another way) means that if we write $\boldsymbol{F}_{\mu\nu} = F^p_{\mu\nu}\boldsymbol{\Sigma}_p$ and assemble

$$\vec{F}_{\mu\nu} \equiv \begin{pmatrix} F^1_{\mu\nu} \\ F^2_{\mu\nu} \\ F^3_{\mu\nu} \end{pmatrix} = (\vec{A}_{\beta,\alpha} - \vec{A}_{\alpha,\beta}) + 2g\vec{A}_\alpha \times \vec{A}_\beta \tag{103}$$

then we have

$$\vec{F}_{\mu\nu} \longrightarrow \vec{F}\,'_{\mu\nu} = \mathbb{R}\vec{F}_{\mu\nu} \tag{102.2}$$

Comparison of (102.1) and (102.2) presents an instance of the well-known connection between $SU(2)$ and $O(3)$.

To kill the subscripts we proceed now in imitation of (16.2), constructing

$$\boldsymbol{F}_{\alpha\beta}\boldsymbol{F}^{\alpha\beta}$$

which is Lorentz invariant, and responds to gauge transformations by similarity transformation. From this result it follows that

- the eigenvalues of the 8×8 matrix $\boldsymbol{F}^2 \equiv \boldsymbol{F}_{\alpha\beta}\boldsymbol{F}^{\alpha\beta}$ are gauge-invariant; equivalently
- the coefficients in $\det(\boldsymbol{F}^2 - \lambda \boldsymbol{I})$ are gauge-invariant; equivalently[42]
- the traces of integral powers of $\boldsymbol{F}^2$ are gauge-invariant.

Proceeding in imitation of our experience in simpler situations (but from no higher necessity[43]) we construct

$$\mathcal{L}_{\text{free gauge field}}(\boldsymbol{A}, \partial\boldsymbol{A}) = \tfrac{1}{4}\text{tr}\big\{\boldsymbol{F}_{\alpha\beta}\boldsymbol{F}^{\alpha\beta}\big\} = \tfrac{1}{4}g^{\alpha\rho}g^{\beta\sigma}F^p_{\alpha\beta}F^q_{\rho\sigma}\,\text{tr}\big\{\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q\big\}$$

But

$$\begin{aligned}
\text{tr}\big\{\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_q\big\} &= \text{tr}\big\{(\boldsymbol{\sigma}_p \otimes \boldsymbol{I}_4)(\boldsymbol{\sigma}_q \otimes \boldsymbol{I}_4)\big\} && \text{by (88)}\\
&= \text{tr}\big\{(\boldsymbol{\sigma}_p\boldsymbol{\sigma}_q \otimes \boldsymbol{I}_4)\big\} && \text{by (85.6)}\\
&= 4\,\text{tr}\big\{\boldsymbol{\sigma}_p\boldsymbol{\sigma}_q\big\} && \text{by (85.5)}\\
&= 8\,\delta_{pq} && \text{by (76)}\\
&= k_{pq} && \text{by (82)}
\end{aligned}$$

So we have[44]

$$\mathcal{L}_{\text{free gauge field}}(\boldsymbol{A}, \partial\boldsymbol{A}) = \tfrac{1}{4}g^{\alpha\rho}g^{\beta\sigma}k_{pq}F^p_{\alpha\beta}F^q_{\rho\sigma} \tag{104}$$

$$\uparrow$$

$$F^p_{\alpha\beta} \equiv A^p_{\beta,\alpha} - A^p_{\alpha,\beta} + gc_u{}^p{}_v A^u_\alpha A^v_\beta \tag{105}$$

and from

$$\left\{\partial_\mu \frac{\partial}{\partial A^r_{\nu,\mu}} - \frac{\partial}{\partial A^r_\nu}\right\}\mathcal{L}_{\text{free gauge field}} = 0$$

---

[42] See p. 13 of "Some applications of an elegant formula due to V. F. Ivanoff" in COLLECTED SEMINARS (1963/70).

[43] It would be interesting on some future occasion to *evaluate* $\det(\boldsymbol{F}^2 - \lambda\boldsymbol{I})$, to see what expressions $\text{tr}(\boldsymbol{F}^{2n})$ actually contribute to the coefficients, and to see whether incorporation of such higher-order terms into the Lagrangian leads to a useful generalization of standard theory... else to identify the principle which forces their exclusion.

[44] It should be noticed that (104) presents not only terms of the type $(\partial A)^2$ first encountered at (16.2), but also terms of the types $A^2\partial A$ and A^4; we can anticipate that *the free gauge field equations will be non-linear*. Notice also that expressions of the design $g^{\alpha\beta}\varkappa^2_{pq}A^p_\alpha A^q_\beta$—analogs of the $g^{\alpha\beta}\varkappa^2 A_\alpha A_\beta$ contemplated earlier—are Lorentz-invariant but *not gauge-invariant; it becomes therefore impossible to assign "mass" to the gauge fields* in any straightforward, gauge-symmetric way.

compute

$$\left.\begin{aligned} \partial_\mu F_r^{\mu\nu} &= \tfrac{1}{C} s_r^\nu \\ s_r^\nu &\equiv c\frac{\partial}{\partial A_\nu^r}\mathfrak{L}_{\text{free gauge field}} = gcF_p^{\nu\alpha} c_r{}^p{}_q A_\alpha^q \end{aligned}\right\} \tag{106.1}$$

of which

$$\left.\begin{aligned} \partial_\mu \vec{F}^{\mu\nu} &= \tfrac{1}{C}\vec{s}^{\,\nu} \\ \vec{s}^{\,\nu} &\equiv 2gc\vec{A}_\alpha \times \vec{F}^{\,\nu\alpha} \end{aligned}\right\} \tag{106.2}$$

provides a more picturesque account (but an account available only within the $SU(2)$ theory, and then only if we have elected to work in the Pauli basis). As yet a third alternative we have this *basis-independent* representation of free motion of the gauge field system:

$$\left.\begin{aligned} \partial_\mu \boldsymbol{F}^{\mu\nu} &= \tfrac{1}{C}\boldsymbol{s}^\nu \\ \boldsymbol{s}^\nu &\equiv igc[\boldsymbol{F}^{\nu\alpha}, \boldsymbol{A}_\alpha] \end{aligned}\right\} \tag{106.3}$$

Equations (106) become meaningful/informative only after one has imported—"by hand," as it were—from (105/3/1) the definition of $F_r^{\mu\nu}$ else $\vec{F}^{\mu\nu}$ else $\boldsymbol{F}^{\mu\nu}$. There is, however, a way to circumvent this formal blemish (if such it be): borrowing a trick from the theory of Procca fields,[45] let us, in place of (104), write

$$\mathfrak{L}_{\text{free gauge}} = \tfrac{1}{2}\big\{F_p^{\alpha\beta}\big[A^p_{\beta,\alpha} - A^p_{\alpha,\beta} + gc_u{}^p{}_v A^u_\alpha A^v_\beta\big] - \tfrac{1}{2}F_p^{\alpha\beta}F^p_{\alpha\beta}\big\} \tag{107}$$

and agree to construe A_p^μ and $F_p^{\mu\nu} = -F_p^{\nu\mu}$ to be *independent* fields. We then have a *pair* of Lagrange equations

$$\left\{\partial_\mu \frac{\partial}{\partial F^r_{\rho\sigma,\mu}} - \frac{\partial}{\partial F^r_{\rho\sigma}}\right\}\mathfrak{L}_{\text{free gauge}} = 0$$
$$\left\{\partial_\mu \frac{\partial}{\partial A^r_{\nu,\mu}} - \frac{\partial}{\partial A^r_\nu}\right\}\mathfrak{L}_{\text{free gauge}} = 0$$

The former yields (105) *as a field equation*, while the later reproduces (106.1).

The free motion of the gauge field system is, according to (106), described by a coupled system of *non-linear* partial differential equations. The system is—owing to the presence of the current-like s-term on the right hand side—"self-excited." That the latter phenomenon is an artifact of non-commutivity (i.e., of the non-Abelian character of the gauge group) is most vividly evident in (106.3).

[45] See again (1–31).

To describe, finally, the dynamics of the full locally $SU(2)$-invariant Dirac theory we assemble

$$\mathcal{L}_2 = \hbar c\Big[\tfrac{1}{2}i\Big\{\tilde{\psi}\boldsymbol{\Gamma}^\alpha\big[\psi_{,\alpha} - igA^p_\alpha\boldsymbol{\Sigma}_p\psi\big] - \big[\tilde{\psi}_{,\alpha} + ig\tilde{\psi}A^p_\alpha\boldsymbol{\Sigma}_p\big]\boldsymbol{\Gamma}^\alpha\psi\Big\} - \varkappa\tilde{\psi}\psi\Big] + \mathcal{L}_{\text{free gauge}} \tag{108}$$

and from

$$\left\{\partial_\mu\frac{\partial}{\partial\tilde{\psi}_{,\mu}} - \frac{\partial}{\partial\tilde{\psi}}\right\}\mathcal{L}_2 = 0 \quad \text{and its adjoint}$$

$$\left\{\partial_\mu\frac{\partial}{\partial F^r_{\rho\sigma,\mu}} - \frac{\partial}{\partial F^r_{\rho\sigma}}\right\}\mathcal{L}_2 = 0$$

$$\left\{\partial_\mu\frac{\partial}{\partial A^r_{\nu,\mu}} - \frac{\partial}{\partial A^r_\nu}\right\}\mathcal{L}_2 = 0$$

obtain (compare (72))

$$\left.\begin{aligned}\big[\boldsymbol{\Gamma}^\mu(\boldsymbol{\partial}_\mu - ig\boldsymbol{A}_\mu) + i\varkappa\boldsymbol{I}\big]\psi = 0 \quad &\text{and its adjoint}\\ F^p_{\mu\nu} = \partial_\mu A^p_\nu - \partial_\nu A^p_\mu + gc_u{}^p{}_v A^u_\mu A^v_\nu \quad &\\ \partial_\mu F^{\mu\nu}_p = \tfrac{1}{C}(J^\nu_p + s^\nu_p) \quad &\end{aligned}\right\} \tag{109}$$

where s^ν_p are the self-interaction currents first encountered at (106), and where the currents J^ν_p can in terms of the fluxes introduced at (91) be described

$$\begin{aligned} J^\nu_p &\equiv eQ^\nu_p \quad \text{with} \quad e \equiv g\hbar c\\ &= ec\tfrac{1}{2}\big\{\tilde{\psi}\boldsymbol{\Gamma}^\nu\boldsymbol{\Sigma}_p\psi + \tilde{\psi}\boldsymbol{\Sigma}_p\boldsymbol{\Gamma}^\nu\psi\big\}\end{aligned} \tag{110}$$

The *global* $SU(2)$-invariance of the Lagrangian $\mathcal{L}_0$ from which we started led to $\partial_\nu J^\nu_p = 0$, but with the adjustment $\mathcal{L}_0 \to \mathcal{L}_2$ those conservation laws have been lost; in their place one has

$$\partial_\nu\mathcal{J}^\nu_p = 0 \quad \text{with} \quad \mathcal{J}^\nu_p \equiv J^\nu_p + s^\nu_p \tag{111.1}$$

which can be read as an immediate consequence of the antisymmetry of $F^{\mu\nu}_p$, and speak to the conservation of

$$\mathcal{J}_p = \int \mathcal{J}^0_p\, dx^1dx^2dx^3 \tag{111.2}$$

At (110) we set $g = \frac{e}{\hbar c}$ to maximize the "electromagnetic appearance" of our results, but abandon any notion that e may have something to do with *electric* charge: e is a new kind of coupling constant—"$SU(2)$-charge." Notice that e serves to describe also the strength of the *self*-interaction, which (as was previously remarked) is a symptom of the non-Abelian character of the gauge group. Gauge theory—conceived by Shaw[8] to be a theory of field interactions—has become now a theory also of intricately structured *self*-interactions.

Dirac theory with local SU(N) gauge invariance. The hard work lies now behind us; we have now only to retrace the argument of the preceding section and to make adjustments at those few points where we drew on properties *specific* to $SU(2)$. We begin (compare (86)) by preparing the canvas; i.e., by assembling the $4N$-component complex field

$$\psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \vdots \\ \psi^N \end{pmatrix} \quad \text{with} \quad \psi^a = \begin{pmatrix} \psi^a_1 \\ \psi^a_2 \\ \psi^a_3 \\ \psi^a_4 \end{pmatrix} \quad : \quad a = 1, 2, \ldots, N \tag{112.1}$$

and by writing

$$\mathcal{L}_0(\psi, \tilde{\psi}, \partial\psi, \partial\tilde{\psi}) = \hbar c \Big[\tfrac{1}{2} i \big\{ \tilde{\psi} \boldsymbol{\Gamma}^\alpha \psi_{,\alpha} - \tilde{\psi}_{,\alpha} \boldsymbol{\Gamma}^\alpha \psi \big\} - \tilde{\psi} \boldsymbol{\varkappa} \psi \Big] \tag{112.2}$$

with

$$\boldsymbol{\Gamma}^\mu \equiv \boldsymbol{I}_N \otimes \boldsymbol{\gamma}^\mu \quad \text{and} \quad \boldsymbol{\varkappa} \equiv \begin{pmatrix} \varkappa_1 & 0 & \ldots & 0 \\ 0 & \varkappa_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \varkappa_N \end{pmatrix} \otimes \boldsymbol{I}_4 \tag{112.3}$$

From the block design of those matrices it follows that the component fields ψ^a are uncoupled in (112.2), which could be written

$$\mathcal{L}_0(\psi, \tilde{\psi}, \partial\psi, \partial\tilde{\psi}; \boldsymbol{\varkappa}) = \sum_{a=1}^{N} \mathcal{L}_0(\psi^a, \tilde{\psi}^a, \partial\psi^a, \partial\tilde{\psi}^a; \varkappa_a)$$

Let $\boldsymbol{s}$ be $N \times N$, unimodular and unitary; the matrices $\boldsymbol{s}$ comprise the "natural elements" of $SU(N)$, while the matrices $\boldsymbol{S} \equiv \boldsymbol{s} \otimes \boldsymbol{I}_4$ give rise to a $4N$-dimensional unimodular unitary representation of $SU(N)$. One can always write

$$\boldsymbol{s} = e^{i\boldsymbol{h}} \quad \text{where } \boldsymbol{h} \text{ is } N \times N \text{ traceless hermitian} \tag{112.1}$$

Write

$$\boldsymbol{h} = h^1 \boldsymbol{\sigma}_1 + h^2 \boldsymbol{\sigma}_2 + \cdots h^{\mathcal{N}} \boldsymbol{\sigma}_{\mathcal{N}} \quad : \quad \mathcal{N} \equiv N^2 - 1 \tag{112.2}$$

where $\big\{ \boldsymbol{\sigma}_p : p = 1, 2, \ldots, \mathcal{N} \big\}$ refer now to an *arbitrary* basis in the vector space $\mathcal{V}_{\mathcal{N}}$ of such (N-dimensional traceless hermitian) matrices. Necessarily there exist real structure constants $c_p{}^r{}_q$ such that

$$[\boldsymbol{\sigma}_p, \boldsymbol{\sigma}_q] = i c_p{}^r{}_q \boldsymbol{\sigma}_r \tag{112.3}$$

and from which we construct the $\mathcal{N} \times \mathcal{N}$ matrices $\mathbb{C}_r \equiv \| i c_r{}^m{}_n \|$ which were seen at (81) to provide the "adjoint representation" of (112.3):

$$[\mathbb{C}_p, \mathbb{C}_q] = i c_p{}^r{}_q \mathbb{C}_r \tag{113}$$

We agree to use

$$k_{pq} = \mathrm{tr}\,\mathbb{C}_p\mathbb{C}_q \quad : \quad \text{elements of } \Bbbk \equiv \|k_{pq}\| \tag{114}$$

and the elements k^{pq} of $\Bbbk^{-1}$ to lower/raise "gauge indices."

Ascend now from N to $4N$ dimensions, it follows straightforwardly from $\boldsymbol{S} \equiv \boldsymbol{s} \otimes \boldsymbol{I}_4$ and properties of the Kronecker product that

$$\boldsymbol{S} = e^{i\boldsymbol{H}} \quad \text{with} \quad \boldsymbol{H} = h^1\boldsymbol{\Sigma}_1 + h^2\boldsymbol{\Sigma}_2 + \cdots h^{\mathcal{N}}\boldsymbol{\Sigma}_{\mathcal{N}} \tag{115.1}$$

where

$$\boldsymbol{\Sigma}_p \equiv \boldsymbol{\sigma}_p \otimes \boldsymbol{I}_4 \quad \text{is } 4N \times 4N \text{ traceless hermitian} \tag{115.2}$$

The structure constants still serve

$$[\boldsymbol{\Sigma}_p, \boldsymbol{\Sigma}_q] = ic_p{}^r{}_q\boldsymbol{\Sigma}_r \tag{115.3}$$

Look now to the response of $\mathcal{L}_0$ to $\psi \longrightarrow \boldsymbol{S}\psi$: we have

$$\tilde{\boldsymbol{S}}\boldsymbol{\Gamma}^\mu\boldsymbol{S} = (\tilde{\boldsymbol{s}} \otimes \boldsymbol{I}_4)(\boldsymbol{I}_N \otimes \boldsymbol{\gamma}^\mu)(\boldsymbol{s} \otimes \boldsymbol{I}_4) = \tilde{\boldsymbol{s}}\boldsymbol{s} \otimes \boldsymbol{\gamma}^\mu = (\boldsymbol{I}_N \otimes \boldsymbol{\gamma}^\mu) = \boldsymbol{\Gamma}^\mu$$

but (by the same line of reasoning)

$$\tilde{\boldsymbol{S}}\boldsymbol{\varkappa}\boldsymbol{S} = \boldsymbol{\varkappa} \quad \text{if and only if all the } \varkappa_a \text{ are set equal}$$

This we do, rendering $\mathcal{L}_0$ "globally $SU(N)$-invariant by design," and Noether hands us a collection of $\mathcal{N} \equiv N^2 - 1$ conservation laws

$$\partial_\mu Q^\mu_r = 0 \text{ with } Q^\mu_r \equiv \tfrac{1}{2}c\tilde{\psi}(\boldsymbol{\Gamma}^\mu\boldsymbol{\Sigma}_r + \boldsymbol{\Sigma}_r\boldsymbol{\Gamma}^\mu)\psi \;\; : \;\; r = 1, 2, \ldots, \mathcal{N} \tag{116.1}$$

which are of a design encountered most recently at (110). We note in passing that the manifest global $U(1)$-symmetry of the theory leads a conservation law of similar design:

$$\partial_\mu Q^\mu_0 = 0 \text{ with } Q^\mu_0 \equiv \tfrac{1}{2}c\tilde{\psi}(\boldsymbol{\Gamma}^\mu\boldsymbol{\Sigma}_0 + \boldsymbol{\Sigma}_0\boldsymbol{\Gamma}^\mu)\psi \tag{116.0}$$

where $\boldsymbol{\Sigma}_0$ is but a fancy way of saying $\boldsymbol{I}_N$.[46]

To achieve invariance under *local* $SU(N)$—of which

$$\begin{aligned} \psi \longrightarrow \psi' &= \boldsymbol{S}(x)\psi \\ \boldsymbol{S}(x) &= e^{ig\boldsymbol{\Omega}(x)} \quad : \quad \boldsymbol{\Omega}(x) = \Omega^r(x)\boldsymbol{\Sigma}_r \text{ traceless hermitian} \end{aligned} \tag{117.1}$$

[46] Note that we are now *not* in position to write equations so explicit as (91), since those reflect special properties of a specific basis (the Pauli basis).

is the initial/defining symptom— we make the "minimal coupling" adjustment

$$\begin{aligned}
&\mathcal{L}_0(\psi,\tilde{\psi},\partial\psi,\partial\tilde{\psi}\,)\\
&\quad\downarrow\\
\mathcal{L}_1 =&\mathcal{L}_0(\psi,\tilde{\psi},\mathcal{D}\psi,\mathcal{D}\tilde{\psi}\,)\\
=&\,\hbar c\Big[\tfrac{1}{2}i\big\{\tilde{\psi}\boldsymbol{\Gamma}^\alpha(\psi_{,\alpha}-ig\boldsymbol{A}_\alpha\psi)-(\tilde{\psi}_{,\alpha}+ig\tilde{\psi}\boldsymbol{A}_\alpha)\boldsymbol{\Gamma}^\alpha\psi\big\}-\tilde{\psi}\boldsymbol{\varkappa}\psi\Big]
\end{aligned}$$

and expand the meaning of (117.1) to include (compare (8) and (93))

$$\boldsymbol{A}_\mu \longrightarrow \boldsymbol{A}'_\mu = \boldsymbol{S}\boldsymbol{A}_\mu\boldsymbol{S}^{-1} + i\,\tfrac{1}{g}\,\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1} \tag{117.2}$$

so as to achieve

$$\mathcal{D}_\mu\psi \longrightarrow (\mathcal{D}_\mu\psi)' = \boldsymbol{S}(\mathcal{D}_\mu\psi) \tag{117.3}$$

"Gauge fields"—the $\mathcal{N}$-tuple of vector fields $A^r_\mu(x)$—enter the discussion when we allow ourselves to write

$$\boldsymbol{A}_\mu(x) = A^r_\mu(x)\boldsymbol{\Sigma}_r \tag{118}$$

and are evidently basis-contingent constructs.

It is as a step preparatory to launching the gauge fields into dynamical motion that—appropriating the intricate argument which led us to (101), an argument which is seen to be not at all specific to $SU(2)$, or to the selection of any specific basis (though with the loss of (98.2) we lose the possibility of using the cross product to express our results, as we did at (103) and (106.2))—we construct

$$\boldsymbol{F}_{\mu\nu} = F^r_{\mu\nu}\boldsymbol{\Sigma}_r \equiv (\partial_\mu\boldsymbol{A}_\nu - \partial_\nu\boldsymbol{A}_\mu) - ig\,(\boldsymbol{A}_\mu\boldsymbol{A}_\nu - \boldsymbol{A}_\nu\boldsymbol{A}_\mu) \tag{119.1}$$

$$F^r_{\mu\nu} = (\partial_\mu A^r_\nu - \partial_\nu A^r_\mu) + gc_p{}^r{}_q A^p_\mu A^q_\nu \tag{119.2}$$

—the point being that $\boldsymbol{F}_{\mu\nu}$ responds to gauge transformation (117.2) by the very simple rule

$$\boldsymbol{F}_{\mu\nu} \longrightarrow \boldsymbol{F}'_{\mu\nu} = \boldsymbol{S}\boldsymbol{F}_{\mu\nu}\boldsymbol{S}^{-1} \tag{120.1}$$

We are placed thus in position to assemble an $\mathcal{L}_2$ formally identical[47] to (108), and obtain coupled field equations formally identical to (109).

In a more complete account of this subject it would become natural at this point to inquire into

- the construction of the *stress-energy tensor* of the gauged theory, and details of energy-momentum trade-off between the ψ-field and the gauge fields;
- Belinfante *symmetrization* of the stress-energy tensor (if required);
- *spin* of the gauge field system;
- motion of the center of mass of the gauge field system.

And, of course, it would be instructive to collect *illustrative solutions* of the field equations. I prefer, however, to look to other matters:

[47] I say "formally identical" because the implicit $\sum_p$ in (108) ran on $\{1,2,3\}$, but runs in the general case on $\{1,2,\ldots,\mathcal{N}\}$. Moreover, the $\boldsymbol{\Sigma}_p$ in (108) are taken to refer specifically to the Pauli basis in $\mathcal{V}_3$, but refer now to an arbitrary basis in $\mathcal{V}_\mathcal{N}$.

The argument which gave (102.2) leads in the more general case to the conclusion that (120.1) can be expressed

$$\vec{F}_{\mu\nu} \longrightarrow \vec{F}'_{\mu\nu} = \mathbb{R}\vec{F}_{\mu\nu} \tag{120.2}$$

where

$$\vec{F}_{\mu\nu} \equiv \begin{pmatrix} F^1_{\mu\nu} \\ F^2_{\mu\nu} \\ \vdots \\ F^{\mathcal{N}}_{\mu\nu} \end{pmatrix} \quad \text{and} \quad \mathbb{R} \equiv e^{ig\mathbb{Z}(x)} \quad \text{with} \quad \mathbb{Z} \equiv \Omega^r \mathbb{C}_r$$

We know from (83.2) that

$$\text{the real matrix } i\mathbb{Z} \text{ is } \Bbbk\text{-antisymmetric: } (i\mathbb{Z})^\intercal = -(i\mathbb{Z})$$

and from this it follows that

$$\text{the real matrix } \mathbb{R} \equiv e^{ig\mathbb{Z}} \text{ is } \Bbbk\text{-orthogonal: } \mathbb{R}^\intercal \Bbbk \mathbb{R} = \Bbbk$$

If, in particular, the basis matrices $\{\mathbb{C}_r\}$ are, by contrived pre-arrangement, *tracewise orthonormal* then $\Bbbk = \mathbb{I}_{\mathcal{N}}$, and $\mathbb{R}$ becomes an element of the rotation group $O(\mathcal{N})$. But

$$\begin{aligned} &SU(N) \text{ is an } \mathcal{N} \equiv (N^2-1)\text{-parameter group, while} \\ &O(\mathcal{N}) \text{ is an } \tfrac{1}{2}\mathcal{N}(\mathcal{N}-1) = \tfrac{1}{2}(N^2-1)(N^2-2)\text{-parameter group} \end{aligned}$$

and from data tabulated below

N	$\mathcal{N} \equiv (N^2-1)$	$\frac{1}{2}\mathcal{N}(\mathcal{N}-1)$
2	3	3
3	8	28
4	15	105
⋮	⋮	⋮

we infer on numerological grounds that the correspondence

$$\boldsymbol{S} = \exp\left\{ig\Omega^p \boldsymbol{\Sigma}_p\right\} \in SU(N) \quad \longleftrightarrow \quad \mathbb{R} = \exp\left\{ig\Omega^p \mathbb{C}_p\right\} \in O(\mathcal{N})$$

which (familiarly) serves to associate elements of $SU(2)$ with elements of $O(3)$, can in more general cases $N > 2$ serve only to associate elements of $SU(N)$ with elements of a *subgroup* of $O(\mathcal{N})$. We note in passing that there do, however,

exist other cases in which $\mathcal{N}(N)$ is at least *triangular*:

$$\begin{aligned}
\mathbf{2}^2-1 &= \quad 3=\Delta(2) \quad \text{where } \Delta(n)\equiv\sum_{k=1}^{n}k \text{ is the } n^{\text{th}} \text{ triangular number}\\
\mathbf{4}^2-1 &= \quad 15=\Delta(5)\\
11^2-1 &= \quad 120=\Delta(15)\\
23^2-1 &= \quad 528=\Delta(32)\\
\mathbf{64}^2-1 &= \quad 4095=\Delta(90)\\
134^2-1 &= \quad 17955=\Delta(189)\\
373^2-1 &= 139128=\Delta(527)\\
&\vdots
\end{aligned}$$

and since $O(n)$ is a $\Delta(n-1)$-parameter group it is at least conceivable that an association of (say) the form $SU(4)\longleftrightarrow O(6)$ is possible. Relatedly, the theory of Clifford algebras inspires interest in numbers of the form 2^p-1, and Ramanujan has observed that in three and only three cases is *such* a number triangular. Each of those cases appears (boldface) in the preceding list; the $p=2$ and $p=4$ are of well-established physical importance (Pauli algebra, Dirac algebra) and it seems to me plausible that the final case $p=12$ may also possess latent physical significance. But it is difficult to manage an algebra with 4095 $\boldsymbol{\Sigma}$-matrices, and my occasional efforts to develop that hunch have thus far been fruitless. Returning now to less speculative matters...

Bringing (120.1) to the field equation $\partial_\mu F_p^{\mu\nu}=\frac{1}{C}\mathcal{J}_p^\nu$ we infer that the conserved net current $\boldsymbol{\mathcal{J}}^\nu\equiv\mathcal{J}^{p\nu}\boldsymbol{\Sigma}_p$ responds to local gauge transformation by the complicated rule

$$\begin{aligned}
\boldsymbol{\mathcal{J}}^\nu\longrightarrow\boldsymbol{\mathcal{J}}'^\nu &= \boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1}\cdot\boldsymbol{F}'^{\mu\nu}+\boldsymbol{S}\boldsymbol{\mathcal{J}}^\nu\boldsymbol{S}^{-1}+\boldsymbol{F}'^{\mu\nu}\cdot\boldsymbol{S}(\boldsymbol{S}^{-1})_{,\mu}\\
&= \boldsymbol{S}\boldsymbol{\mathcal{J}}^\nu\boldsymbol{S}^{-1}+[\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1},\boldsymbol{F}'^{\mu\nu}]
\end{aligned}\tag{121.1}$$

Its response to *global* gauge transformation is, however, simple: the commutator drops away (because $\boldsymbol{S}_{,\mu}=\mathbf{0}$), leaving

$$\boldsymbol{\mathcal{J}}^\nu\longrightarrow\boldsymbol{\mathcal{J}}'^\nu=\boldsymbol{S}\boldsymbol{\mathcal{J}}^\nu\boldsymbol{S}^{-1}\tag{121.21}$$

which can be written

$$\vec{\mathcal{J}}^{\,\nu}\longrightarrow\vec{\mathcal{J}}'^{\nu}=\mathbb{R}\,\vec{\mathcal{J}}^{\,\nu}\tag{121.22}$$

It was this circumstance (together with the circumstance that in $SU(2)$ theory the 3×3 matrix $\mathbb{R}\in O(3)$) which led Yang & Mills to the satisfying conclusion that "total isotopic spin" $\vec{\mathcal{J}}\equiv\int\vec{\mathcal{J}}^0dx^1dx^2dx^3$ is a *vector*, which responds to (global) gauge transformation by "rotation in isotopic spin space."

It is (recall (110)) through

$$\boldsymbol{J}=ec\tfrac{1}{2}\{\tilde{\psi}\boldsymbol{\Gamma}^\nu\boldsymbol{\Sigma}^p\psi+\tilde{\psi}\boldsymbol{\Sigma}^p\boldsymbol{\Gamma}^\nu\psi\}\boldsymbol{\Sigma}_p\tag{122}$$

that the gauge fields sense the presence of the Dirac field ψ. The right side of the preceding equation is easily seen to be basis-independent, but I have been unable to discover any natural way to formulate (122) which does not make incidental *reference* to a basis... which strikes me as curious.

General observations, and some topics which might be included in a more comprehensive account of gauge field theory. In the preceding discussion we took "stacked copies of the Dirac equation" as our point of departure. We could instead have taken "stacked copies of the Klein-Gordon equation" or "stacked copies of the Procca equation"... and—particularly if we worked in canonical formalism—would be led to results formally identical (or nearly so) to the results now in hand (though the specific meaning of the $\boldsymbol{\Gamma}$-matrices would vary from case to case).[48]

Elementary calculus supplies the statement

$$(\partial_\mu\partial_\nu - \partial_\nu\partial_\mu)(\text{any nice function}) = 0$$

which we abbreviate $(\partial_\mu\partial_\nu - \partial_\nu\partial_\mu) = 0$. But from the definition (10)

$$\mathcal{D}_\mu \equiv \partial_\mu - igA_\mu(x)$$

it follows on the other hand that

$$\mathcal{D}_\mu\mathcal{D}_\nu - \mathcal{D}_\nu\mathcal{D}_\mu = -ig(A_{\nu,\mu} - A_{\mu,\nu})$$

while in the non-Abelian case

$$\boldsymbol{\mathcal{D}}_\mu \equiv \boldsymbol{\partial}_\mu - ig\boldsymbol{A}_\mu(x)$$

we obtain

$$\begin{aligned}\boldsymbol{\mathcal{D}}_\mu\boldsymbol{\mathcal{D}}_\nu - \boldsymbol{\mathcal{D}}_\nu\boldsymbol{\mathcal{D}}_\mu &= -ig\big\{(\boldsymbol{A}_{\nu,\mu} - \boldsymbol{A}_{\mu,\nu}) - ig[\boldsymbol{A}_\mu, \boldsymbol{A}_\nu]\big\} \\ &= -ig\,\boldsymbol{F}_{\mu\nu} \quad \text{by (119.1)}\end{aligned} \tag{123}$$

We are not yet in position to comment on the deeper significance of this striking result, except to remark that it makes transparently clear how $\boldsymbol{F}_{\mu\nu}$ acquired its especially simple gauge transformation properties: it inherited them from $\boldsymbol{\mathcal{D}}_\mu$.

In Maxwellian electrodynamics we learn that it is from the sourceless equations

$$\boldsymbol{\nabla}\cdot\boldsymbol{B} = 0 \quad \text{and} \quad \boldsymbol{\nabla}\times\boldsymbol{E} + \partial_0\boldsymbol{B} = \boldsymbol{0} \tag{124.1}$$

—which is to say: from

$$\partial_\mu G^{\mu\nu} = 0 \quad \text{with} \quad G_{\mu\nu} \equiv \tfrac{1}{2}\epsilon_{\mu\nu\alpha\beta}F^{\alpha\beta} \tag{124.2}$$

which can be expressed alternatively as a quartet of "windmill sum" relations

$$\epsilon^{\mu\nu\alpha\beta}\partial_\nu F_{\alpha\beta} = 0 \tag{124.3}$$

—that we acquire license to write

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{125}$$

[48] Could we proceed similarly from "stacked copies of the Hamilton-Jacobi equation" or "stacked copies of the relativistic free particle equation," and thus produce non-Abelian generalizations of the theories developed on pp. 8–20 and pp. 20–26? Multi-component field systems are commonplace in field theory, but what might be the physical interpretation of "stacked copies of the classical mechanics of a particle"? Could such formalism be associated with the classical physics of particles with internal degrees of freedom ("spin")? These are questions to which I hope to return on another occasion.

And that when one *does* write (125) then equations (124) become *automatic*. In non-Abelian gauge field theory ("generalized electrodynamics") one has only to introduce (123) into the following instance

$$[\boldsymbol{\mathcal{D}}_\nu, [\boldsymbol{\mathcal{D}}_\alpha, \boldsymbol{\mathcal{D}}_\beta]] + [\boldsymbol{\mathcal{D}}_\alpha, [\boldsymbol{\mathcal{D}}_\beta, \boldsymbol{\mathcal{D}}_\nu]] + [\boldsymbol{\mathcal{D}}_\beta, [\boldsymbol{\mathcal{D}}_\nu, \boldsymbol{\mathcal{D}}_\alpha]] = \mathbf{0}$$

of Jacobi's identity to obtain

$$[\boldsymbol{\mathcal{D}}_\nu, \boldsymbol{F}_{\alpha\beta}] + [\boldsymbol{\mathcal{D}}_\alpha, \boldsymbol{F}_{\beta\nu}] + [\boldsymbol{\mathcal{D}}_\beta, \boldsymbol{F}_{\nu\alpha}] = \mathbf{0} \tag{126}$$

or (more compactly)

$$\epsilon^{\mu\nu\alpha\beta}[\boldsymbol{\mathcal{D}}_\nu, \boldsymbol{F}_{\alpha\beta}] = \epsilon^{\mu\nu\alpha\beta}[\boldsymbol{\partial}_\nu, \boldsymbol{F}_{\alpha\beta}] - ig\epsilon^{\mu\nu\alpha\beta}[\boldsymbol{A}_\nu, \boldsymbol{F}_{\alpha\beta}] = \mathbf{0}$$

In the Abelian (i.e., in the Maxwellian) case the second term on the right drops away, and the surviving first term, when allowed to operate on the function 1, gives back (124.3). It will appreciated that equations (126) are not field equations but *identities*, satisfied in every instance and necessarily by $\boldsymbol{F}_{\mu\nu}$ (which is to say: by $\boldsymbol{A}_\mu$) in consequence of the manner in which those objects were defined. Equations (126) play within gauge field theory a role analogous (and abstractly identical) to the roll played by the so-called "Bianchi identities" in general relativity.[49]

It can be argued that the exterior calculus provides the language of choice for developing formal (and some computational) properties of the theory of Abelian gauge fields (Maxwellian electrodynamics).[50] The question arises: can a "generalized exterior calculus" be devised which serves equally well to illuminate the essentials of non-Abelian gauge field theory?

Maxwellian electrodynamics is well-known to be invariant with respect to "duality rotation"—an internal symmetry of which

$$\begin{aligned}\boldsymbol{E} &\longrightarrow \boldsymbol{E}' = \boldsymbol{E}\cos\theta + \boldsymbol{B}\sin\theta\\ \boldsymbol{B} &\longrightarrow \boldsymbol{B}' = \boldsymbol{B}\cos\theta - \boldsymbol{E}\sin\theta\end{aligned}$$

captures the simplified essence.[51] Can a similar symmetry be identified in the non-Abelian case?

It was a $U(1) \otimes SU(2)$ gauge theory which Glashow-Weinberg and Salam used in the late 1960's to achieve unification of the electromagnetic and weak interactions (see Quigg's Chapter 6), a $SU(3)$ gauge theory which (at about the

[49] See Michio Kaku, *Quantum Field Theory: A Modern Introduction* (1993), p. 297; M. E. Peskin & D. V. Schroeder, *An Introduction to Quantum Field Theory* (1995), p. 500 and/or the index of any good general relativity text.

[50] See my "Electrodynamical applications of the exterior calculus," (1996).

[51] See §7 in the material just cited; also pp. 327–331 in ELECTRODYNAMICS (1972) and p. 51 in Chris Quigg's *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions* (1983).

same time) called colored quarks into being and resulted in the creation of QCD (quantum chromodynamics; see Quigg's Chapter 8), and the so-called Standard Model—which unites the whole shebang and accounts satisfactorally for most of the observational evidence—is a $U(1) \otimes SU(2) \otimes SU(3)$ gauge theory. It would be interesting—but a major undertaking—to trace the classical outlines[52] of that work. One would come away from such an exercise with a sense of how one finds "wiggle-room" within the fairly rigid framework provided by the gauge field idea... and of what contortions Nature herself appears to require; more particularly, one would acquire a sense of how difficult it is to endow gauge fields with mass, and of how wonderfully ingenious (if in some respects still unsatisfactory) has been the effort to do so (Quigg's Chapter 5). But for discussion of those topics I must—for now—be content to refer my readers to the abundant literature.[53]

[52] By "classical outlines" I mean "up to the point of quantization." It is, of course, quantization which lends physical significance the theory. But it opens a can of mathematical worms which have no place in an account of the elements of *classical* field theory.

[53] I have made references to Quigg,[51] who is often especially clear, and supplies good bibliographic information, but one should also consult Chapter 11 in Griffiths[4] and relevant paragraphs in (say) Kaku and Peskin & Schroeder[49]. The literature is, as I say, vast; for a random taste of its riches see the essay "Secret symmetry: an introduction to spontaneous symmetry breakdown and gauge fields" in S. Coleman, *Aspects of Symmetry* (1985).

4

GRAVITATIONAL FIELD THEORY

Introduction. It is by "instantaneous action at a distance"—more specifically: by exerting attractive central forces of strength

$$F = Gm_1m_2/r^2 \tag{1}$$

upon one another—that, according to Newton (*Principia Mathematica* 1686), bodies interact gravitationally. Building upon this notion (to which many of his continental contemporaries objected on *philosophical* grounds), he was able to account theoretically for Kepler's emperical "laws of planetary motion" (1610) and to lay the foundation for a famously successful celestial mechanics.[1]

Special Relativity (1905) declared Newtonian dynamics to be defective because not Lorentz covariant, and Newton's Law of Universal Gravitation to be untenable because it drew upon a concept—"distant simultaneity"—which relativity had rendered obsolete. It placed upon physicists the burden of devising a "relativistic theory of gravitation"... not to account for some disagreement between Newtonian theory and observation (of which, if we set aside a little problem concerning the precession of Mercury's orbit, there were none), but to achieve logical consistency.

[1] "Celestial mechanics" is an archaic term for what we would today call "planetary mechanics," and contains an echo of Newton's question (fairly radical for the time): Does gravity—the gravity which causes apples to fall—extend all the way to the moon? Is the moon "falling" around the earth? For a good account of this ancient history (the details of which are much more convoluted/interesting than one might imagine) see G. E. Christenson, *This Wild Abyss: The Story of the Men Who Made Modern Astronomy* (1978).

At p. 1265 in the index of Misner, Thorne & Wheeler's *Gravitation*[2], under the heading "Gravitation, theories of," one encounters mention of (among others)

- Bergmann's theory
- Cartan's theory
- Coleman's theory
- Kustaanheimo's theory
- Ni's theory
- Nordstrøm's theory
- Papapetrou's theory
- Whitehead's theory
- the flat spacetime theories of Gupta, Kraichnan, Thirring, Feynman, Weinberg, Deser and others

and at p. 1049 a description of a "parameterized post-Newtonian formalism" which was, as an aid to experimentalists, developed in the 1970's to provide *simultaneous* expression of most of the theories listed above... each of which was originally put forward in response to the logical consistency problem just mentioned. Consulting the list of Einstein's publications[3] we find that he himself first addressed the problem in 1907; a second paper appeared in 1911, five more in 1912, and those (though he remained intensely involved in a variety of other problems) were followed by a flood of gravitational papers up until the publication, in 1915, of an account of "general relativity" in its finished form.

Einstein's own point of departure was provided by what he called "the happiest thought of my life."[4] It was November of 1907 and Einstein was, as he later wrote,

> *"... sitting in a chair in the patent office at Bern when all of a sudden a thought occurred to me: 'If a person falls freely he will not feel his own wwight.' I was startled. This simple thought made a deep impression upon me. It impelled me toward a theory of gravitation."*

Thus did the Principle of Equivalence spring into being. It holds "gravitational force" and "the fictitious force which arises from acceleration relative to the local inertial frame(s)"—in short: it holds gravitation and non-inertiality—to be physically indistinguishable; i.e., to be different names for the same thing. The Principle of Equivalence contributed little or not at all to most of the early efforts to construct a "relativistic theory of gravitation," but exerted a

---

[2] I will have frequent occasion to refer to this "Black Bible" of gravitation theory, which appeared in 1973 but remains in many respects definitive. I will, on such occasions, use the abbreviation *MTW*.

[3] Such a list—not quite complete—can be found on pp. 689–760 in Paul Schlipp's *Albert Einstein: Philosopher-Scientist* (1951). See also C. Lanczos, *The Einstein Decade (1905–1915)* (1974) and A. Pais, *'Subtle is the Lord...': The Science and the Life of Albert Einstein* (1982) for annotated bibliographic information.

[4] "...die Glücklichste Gedanke meines Lebens;" see Chapter 9 in Pais.[3]

powerful guiding influence upon Einstein's own thought.[5] It was, however, not immediately evident just *how* his happy thought was to be folded into a fully developed theory of gravitation; the journey from special to general relativity took nearly a decade to complete, and was marked by many hesitations, retreats, amendations. The voyage reached its end on 25 November 1915, when Einstein submitted to the Prussian Academy a paper which presented the gravitational field equations in their final form. Five days previously, David Hilbert had submitted to Gesellschaft der Wissenschaften in Göttingen a manuscript containing identical equations.[6]

Field-theoretic aspects of Newton's theory of gravitation. Newton's theory was presented as a theory of 2-body interaction. But it can, by importation of concepts and methods borrowed from electrostatics, readily be portrayed as a rudimentary field theory, and it is in that form that it is most conveniently compared to the full-blown field theories which would supplant it, and from which it must (for well-established observational reasons) be recoverable as the leading approximation.

Let the density function $\rho(\boldsymbol{\xi})$ describe (relative to an inertial frame) the instantaneous distribution of "gravitating matter," and let a test particle of mass (which is to say: of "gravitational charge") m reside momentarily at $\boldsymbol{x}$. To describe the force experienced by the test particle write

$$\boldsymbol{F} = m\boldsymbol{g}(\boldsymbol{x}) \tag{2}$$

[5] It accounts, in particular, for the circumstance that his second gravitational paper (1911) bore the title "Bemerkung zu dem Gesetz von Eötvös." The "Eötvös experiments" (1889 and 1922) looked to the relationship of "inertial mass" to "gravitational mass (or charge)" and established that the ratio

$$\frac{\text{gravitational mass}}{\text{inertial mass}}$$

is "universal" in the sense that it does not vary from material to material by more than 5 parts in 10^9. In the 1960's Robert Dicke used more modern techniques to establish that departures from the Principle of Equivalence cannot exceed one part in 10^{11}. See *MTW* §38.3 for more detailed discussion.

[6] See Pais,[3] §14d. Hilbert, who considered physics to be "too difficult for physicists," imagined himself to be constructing an axiomatic theory of the world (an ambition which Einstein considered to be "too great an audacity... since there are still so many things which we cannot yet remotely anticipate"), and in his grandly titled "Die Grundlagen der Physik" imagined that he had achieved a unified theory of gravitation and electromagnetism. Hilbert's theory is distinguished most importantly from Einstein's by the more prominent role which he assigned to variational principles; we recall that he had retained Emmy Noether to assist him in this work, and it is Noether whom in 1924 he credited for some of his paper's most distinctive details.

Evidently $\boldsymbol{g}(\boldsymbol{x})$ is the gravitational analog of an electrostatic $\boldsymbol{E}$-field. The force-law proposed by Newton is conservative, so $\boldsymbol{\nabla}\times\boldsymbol{g}=\boldsymbol{0}$, from which follows the possibility of writing

$$\begin{aligned}\boldsymbol{g}&=-\boldsymbol{\nabla}\varphi\\ [\varphi]&=(\text{velocity})^2\end{aligned}\tag{3}$$

In mimicry of the electrostatic equation

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}=\rho \quad : \quad \left\{\begin{array}{l}\text{charge density regulates the divergence}\\ \text{of the electrostatic field}\end{array}\right.$$

we write

$$\boldsymbol{\nabla}\cdot\boldsymbol{g}=-4\pi G\rho \quad : \quad \left\{\begin{array}{l}\text{mass density regulates the convergence}\\ \text{of the gravitostatic field}\end{array}\right.\tag{4}$$

where the minus sign reflects the fact that the gravitational interaction is *attractive*, and the 4π was inflicted upon us when Newton neglected to install a $\frac{1}{4\pi}$ in (1). Introducing (3) into (4) we have the gravitational Poisson equation

$$\nabla^2\varphi=4\pi G\rho\tag{5}$$

which in integral formulation reads

$$\begin{aligned}4\pi G\iiint_{\mathcal{R}}\rho\,dx^1dx^2dx^3&=\text{total mass interior to }\mathcal{R}\\ &=\iiint_{\mathcal{R}}\nabla^2\varphi\,dx^1dx^2dx^3\\ &=-\iiint_{\mathcal{R}}\boldsymbol{\nabla}\cdot\boldsymbol{g}\,dx^1dx^2dx^3\\ &=-\iint_{\partial\mathcal{R}}\boldsymbol{g}\cdot\boldsymbol{dS}\\ &=\text{gravitational influx through }\partial\mathcal{R}\end{aligned}\tag{5}$$

Take $\mathcal{R}$ to be, in particular, a sphere centered on a point mass M; then (5) gives $4\pi r^2g(r)=4\pi GM$ whence $g(r)=GM/r^2$ and we have

$$\left.\begin{aligned}\boldsymbol{g}(r)=-GM\,\hat{\boldsymbol{r}}/r^2&=-\boldsymbol{\nabla}\varphi(r)\\ \varphi(r)&=-GM/r\end{aligned}\right\}\tag{6}$$

which describe the gravitational field of an isolated point mass. For a distributed source we have

$$\varphi(\boldsymbol{x})=-G\iiint\{\rho(\boldsymbol{\xi})/|\boldsymbol{x}-\boldsymbol{\xi}|\}\,d\xi^1d\xi^2d\xi^3\tag{7.1}$$

which gives back (6) in the case $\rho(\boldsymbol{\xi})=M\delta(\boldsymbol{\xi})$.

The distinction between "gravitostatics" and "gravitodynamics" did not exist for Newton, since he considered gravitational effects to propagate instantaneously. To describe the gravitational potential engendered by a *moving* mass distribution he would have written

$$\varphi(\boldsymbol{x},t) = -G\iiint \{\rho(\boldsymbol{\xi},t)/|\boldsymbol{x}-\boldsymbol{\xi}|\}\, d\xi^1 d\xi^2 d\xi^3 \tag{7.2}$$

which is to say: he would simply have repeated (7.1) at each incremented value of t. Such a program makes no provision for the "retardation" effects which distinguish electrodynamics from electrostatics.

To describe the motion of a mass point m in the presence of such an imposed field, Newton writes

$$m\ddot{\boldsymbol{x}} = -m\boldsymbol{\nabla}\varphi(\boldsymbol{x},t) \tag{8}$$

from which the m-factors (an inertial mass on the left, a gravitational charge on the right) drop away.

To complete the theory Newton would be obligated by his own 3rd Law to describe the action of m back upon $\rho(\boldsymbol{x},t)$—else to argue that it can, in the specific instance, be neglected—and, more generally, to construct (borrow from fluid dynamics?) a field equation descriptive of $\rho(\boldsymbol{x},t)$; the latter assignment presents one with a continuous analog of the gravitational n-body problem ... where the "problem" is not to write but to *solve* the equations of motion.

Special relativistic generalizations of Newtonian gravitation. If we look upon the potential φ as the object most characteristic of Newtonian gravity—i.e., if we imagine ourselves to be looking for a *relativistic scalar field theory* which gives back Newton's theory in the non-relativistic limit, then it becomes natural in place of (5) to write

$$\left\{\left(\tfrac{1}{c}\tfrac{\partial}{\partial t}\right)^2 - \nabla^2\right\}\varphi(x) = -4\pi G\rho(x) \quad ;i.e. \quad \Box\varphi = -4\pi G\rho$$

since this familiar equation is manifestly covariant, and gives back (5) in the limit $c \uparrow \infty$. In place of (7.2) one would then obtain

$$\varphi(x) = -G\iiiint D_{\scriptscriptstyle R}(x-\xi)\rho(\xi)d\xi^0 d\xi^1 d\xi^2 d\xi^3$$

where the retarded Green's function $D_{\scriptscriptstyle R}(x-\xi)$ vanishes except on the lightcone which extends backward from x.[7]

Relativistic generalization of (8) is more interesting because a bit less straightforward. Notice first that we can *not* simply write $ma^\mu = -m\partial^\mu\varphi$

[7] See ELECTRODYNAMICS (1980), pp. 379–389 for details.

because the Minkowski force on the right is not velocity-dependent, therefore cannot satisfy $K \perp u$, as required.[8] We are led thus to write

$$ma^\mu = K^\mu \quad \text{with (tentatively)} \quad K_\mu = m(\partial_\alpha \varphi)[\delta^\alpha{}_\mu - u^\alpha u_\mu / c^2]$$

because (i) the proposed K_μ depends linearly upon the derivatives of φ and (ii) clearly does yield $K_\mu u^\mu = 0$. Noticing that we have

$$\begin{aligned} K_\mu &= m\{\varphi_{,\mu} - (1/c^2)u_\mu \tfrac{d}{d\tau}\varphi\} \\ &= m\{\varphi_{,\mu} - (1/c^2)\tfrac{d}{d\tau}(u_\mu \varphi) + (1/c^2)\varphi a_\mu\} \end{aligned}$$

and that the final term on the right (since it is itself normal to u) could be abandoned without compromising the normality of what remains . . . we do so, obtaining a refined equation of motion

$$\tfrac{d}{d\tau}\big[m(1 + \varphi/c^2)u^\mu\big] = m\partial^\mu \varphi$$

which can be notated

$$\gamma \tfrac{d}{dt}\big[m(1 + \varphi/c^2)\gamma \begin{pmatrix} c \\ \boldsymbol{v} \end{pmatrix}\big] = m \begin{pmatrix} \frac{1}{c}\partial_t \varphi \\ -\boldsymbol{\nabla}\varphi \end{pmatrix}$$

The spatial part of the preceding equation gives back (8) in the non-relativistic limit $c \uparrow \infty$, which was the point of the exercise. The form of the equation makes it natural to introduce

$$m^* \equiv m(1 + \varphi/c^2) \equiv \text{effective inertial mass}$$

It is interesting that in this theory "gravitational charge" m is invariable, while the "effective inertial mass" is environmentally contingent, and that the two do *not* cancel each other out.[9]

If, on the other hand, we look upon the gravitational 3-vector $\boldsymbol{g}$ as the object most characteristic of Newtonian gravity then it becomes natural to suppose that it possesses a heretofore unnoticed companion $\boldsymbol{h}$—a gravitational analog of magnetism—and to proceed in direct imitation of Maxwellian electrodynamics to an "antisymmetric tensor theory of relativistic gravitation."

These and a number of other purported "special relativistic generalizations of Newtonian gravitation" are discussed in Chapter 7 of *MTW*. All require formal flights of fancy which take one away from the secure observational base of the theory, all at one point or another are contradicted by the observational facts, and some have been found to be internally inconsistent. All, that is, except some of the most recent, which were found—somewhat surprisingly—to be round-the-bush *reconstructions of Einstein's general relativity*, and not the intended alternatives to it. I suspect that Einstein himself considered all such efforts misguided for the simple reason that, in sanctifying special relativity, they created no place at the table for the Principle of Equivalence—no place for what he sometimes called the "relativity of non-uniform motion."

---

[8] See again the discussion preliminary to (3–51).

[9] For further discussion see RELATIVISTIC DYNAMICS (1967), pp. 17–20.

Theoretical program evidently implicit in the Principle of Equivalence. Here in this laboratory we erect a Cartesian $\boldsymbol{y}$-frame, which we naively take to be an inertial frame, with 3-axis pointing up. We have turned off all force fields except gravity

$$\boldsymbol{g} = \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix}$$

(for which we could find no switch, no shielding... else we would have turned if off too). At $t = 0$ we launch pellets in all directions, with all velocities. Each

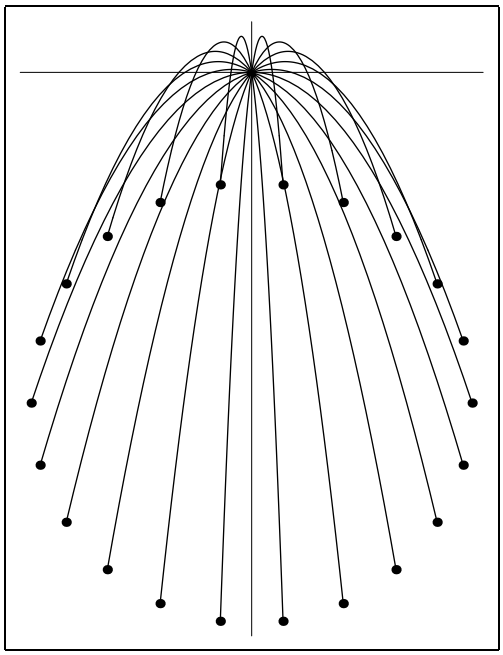


FIGURE 1: *Pellets are launched (all with the same initial speed) in various directions at $t = 0$, and trace parabolic arcs as they fall in the uniform gravitational field. In the figure they have been arrested at the same instant, and are seen to lie on a circle. The envelope of the family of trajectories appears to be parabolic, with the origin at the focus.*

pellet traces a parabolic arc, and the whole display looks like the 4^{th} of July. To describe any particular pellet we write $m\ddot{\boldsymbol{y}} = m\boldsymbol{g}$ and obtain $\boldsymbol{y}(t) = \boldsymbol{v}t + \frac{1}{2}\boldsymbol{g}t^2$, of which

$$\begin{pmatrix} y^1(t) \\ y^2(t) \\ y^3(t) \end{pmatrix} = \begin{pmatrix} v^1 t \\ v^2 t \\ v^3 t - \frac{1}{2}gt^2 \end{pmatrix}$$

provides a more explicit rendition. Now introduce the Cartesian $\boldsymbol{x}$-frame of an observer who (irrotationally) *drops* from the origin at the moment of the

explosion. To describe the time-dependent relationship between the $\boldsymbol{y}$-frame and the $\boldsymbol{x}$-frame we write $\boldsymbol{y} = \boldsymbol{x} + \frac{1}{2}\boldsymbol{g}t^2$ and from $m\ddot{\boldsymbol{y}} = m(\ddot{\boldsymbol{x}} + \boldsymbol{g}) = m\boldsymbol{g}$ conclude that the falling observer writes $m\ddot{\boldsymbol{x}} = \boldsymbol{0}$ to describe the motion of the pellets, which he sees to be moving uniformly and rectilinearly (which is to say: in accordance with Newton's 1st Law): $\boldsymbol{x}(t) = \boldsymbol{v}t$. The relationship between our view of the display and the view presented to the falling observer is illustrated in Figure 2.

Einstein argues that it is the observer in (irrotational) free fall who is the *inertial* observer in this discussion; that we in the laboratory have had to

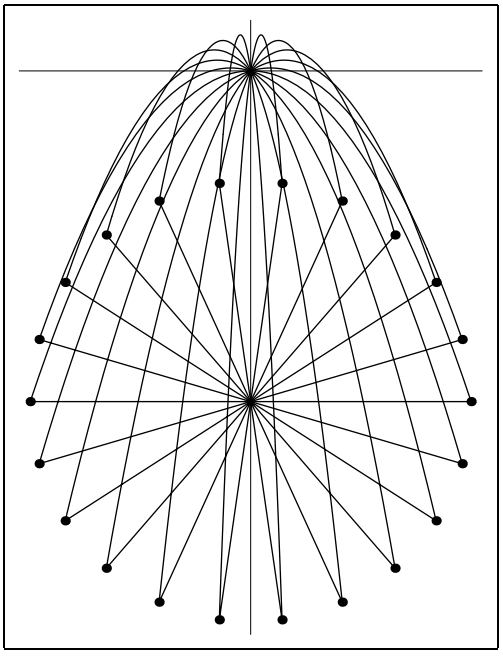


FIGURE 2: *A falling observer experiences no gravitational field, and sees the pellets to be in radial recession. Knowing them to have been launched with the same speed, he is not surprised to observe that they lie on a uniformly expanding circle (sphere).*

"invent gravity" in order to compensate for the circumstance that, relative to the inertial observer, we are accelerating upwards with acceleration g, and that therefore we should not be surprised when it is "discovered" that what we misguidedly call "gravitational charge" is proportional through a universal factor to inertial mass.[10]

[10] Einstein's viewpoint is nicely developed in §2 of *Spacetime Physics*, by Edwin Taylor & John Wheeler (1963). It is advanced on somewhat different grounds in Chapter 1 ("Ground to Stand on: Inertiality & Newton's First Law") of CLASSICAL MECHANICS (1983).

The perception of uniform/rectilinear pellet motion—i.e., of the absence of a gravitational field—would be shared also by other observers $\mathcal{O}', \mathcal{O}'', \ldots$ who are in states of (irrotational) unaccelerated motion with respect to our falling observer $\mathcal{O}$. It is this *population* of observers which in Newtonian physics is interlinked by Galilean transformations, and in relativistic physics by Lorentz transformations.

We nod indulgently at $\mathcal{O}$'s account of events, then observe that "Of course, you will see pellets (and the observers who ride them) to move uniformly/rectilinearly only for awhile—only until they have ventured far enough away to sense non-uniformity of the gravitational field, by which time tidal effects will have begun to distort their spherical pattern."[11] We might write

$$\begin{aligned}\boldsymbol{g}(\boldsymbol{x}) &= \boldsymbol{g}(\mathbf{0}) + \boldsymbol{g}_i(\mathbf{0})x^i + \tfrac{1}{2}\boldsymbol{g}_{ij}(\mathbf{0})x^ix^j + \cdots \\ &= \boldsymbol{g} + \text{tidal terms}\end{aligned}$$

to underscore the force of our remark.

$\mathcal{O}$'s—Einstein's—response to our remark must necessarily be radical, for he cannot reasonably enter into discussion of the higher order properties of something he has already declared to be a delusion. Einstein's embrace of the Principle of Equivalence leads him...

- to borrow from Newton the idea that inertial motion is the motion that results when all forces/interactions that can in principle be turned off/shielded have been;[12]

[11] What we have initially in mind is simply that the "down vector" here is not parallel to the "down vector" over there, but that remark is subject to interesting refinement. Assume the earth to be a (non-rotating) homogeneous sphere of radius R and mass M. The pellets then move in an attractive central force of strength

$$g(r) = \begin{cases} GM\,r^{-2} & \text{if } r \geqslant R \\ (GM/R)\,r^{+1} & \text{if } r \leqslant R \end{cases}$$

The "parabolic arcs" mentioned previously are really sections of Keplerean ellipses (or hyperbolæ, if v is sufficiently great), with the earth's center at one focus. While passing through the earth (which pellets do because we have "turned off all force fields," and "observers" find easy to do because they are mythical) the pellets move as though attached to an isotropic spring, and trace a section of an ellipse with *center* at the center of the earth.

[12] This simple-seeming thought rests upon some heavy idealization. It is one thing to *imagine* turning off the phenomenological forces which would impede a pellet's passage through the earth, but how in practical fact would one turn off the fundamental interactions which underlie those phenomenological forces? We are in something like the predicament of the classical physicist who finds it convenient to "turn off quantum mechanics," and is forced to pretend that he has not thereby precluded the existence of meter sticks.

- to recognize that "gravitational forces" *cannot* be turned off/shielded (can only be *transformed* away); at this point "inertial motion" has become synomymous with what Newton called "free fall;"

- to recognize that inertial observers can engage in special relativistic dialog only so long as—fleetingly—they share the same spatio-temporal "neighborhood," where the bounds of neighborhood are breached when the relative motion of $\mathcal{O}$ and $\mathcal{O}'$ is no longer uniform/rectilinear; in the absence of gravitation all neighborhoods would be co-extensive and infinite (i.e., there would be only one neighborhood, and we could dispense with the concept), but in the presence of gravitation they become local... like little platelets tangent to a curved surface.

Newton's "rectilinear" can be phrased "geodesic" in Euclidean 3-space. His "uniform rectilinear" could be similarly phrased if 4-dimensional spacetime were suitably metrized, and (as Minkowski was the first to emphasize) it was such a "suitable metrization of spacetime" which lay at the heart of Einstein's accomplishment when he invented special relativity. It became therefore natural for Einstein

- to associate the inertial motion of free-falling test particles with *geodesics in a spacetime of suitably altered geometry*.

Newton's assertion that "masses cause one another to depart from inertiality by exerting gravitational forces upon one another" becomes, from this point of view, an assertion that masses cause *no* "departure from inertiality," but instead *alter the geometry* of the spacetime upon which their respective inertial geodesics are inscribed. But Einstein had already established the equivalence of mass and energy. It became therefore natural for him

- to anticipate that the geometry of spacetime is conditioned by the distribution of mass/energy (which is itself in free-fall controlled by the geometry: the world has thus become "self-interactive geometry").[13]

Prior to (and well into) 1912 Einstein had concentrated on *scalar* theories of gravitation. He had achieved what he considered to be good results in the static theory, but was finding the dynamical theory to be "devilishly difficult." On August 10 Einstein registered as a resident of Zürich, to which he had, upon the invitation of Marcel Grossmann, moved from Prague in order to accept an

[13] I must emphasize that it is *as a rhetorical device*—the better to clarify my expository intent—that I have allowed myself to impute motivations to Einstein for which I can, in some instances, provide no specific documentation. My remarks are not (!) intended to be read as "encapsulated history of science." Made-up history is the worst kind of history, and always does violence to the entangled facts. In the present instance it would be a mistake to lose sight of the fact that it took Einstein *several years* to create general relativity, that during those years his motivation was marked by frequent shifts and turns, and that it was not entirely clear where he was headed until he got there.

appointment to the faculty of the ETH. Grossman (1878–1936), it is invariably remarked, had loaned class notes to Einstein when both were students at the ETH. His father had helped Einstein gain employment at the patent office in Bern, while he himself had gone on to become a professor of geometry and (recently) dean of the mathematics–physics section at the ETH. It was, according to Pais,[14] sometime between August 10 and August 16 that Einstein pleaded "Grossmann, Du musst mir helfen, sonst werd' ich verrückt!"[15] and was made aware for the first time of Riemannian geometry, and of the tensor analysis of Ricci and Levi-Civita. Whereupon Einstein recalled that he had, in fact, already been exposed to the Gaussian theory of surfaces in the classroom on one Geisler (whose successor at the ETH was Weyl).[16]

Thus it came about that the scalar gravitation of a Saturday had become a tensor theory by the next Friday. In October of 1912 Einstein wrote to Sommerfeld that

> *"At present I occupy myself exclusively with the problem of gravitation and now believe that I shall master all difficulties with the help of a friendly mathematician here [Grassmann]. But one thing is certain: in all my life I have labored not nearly as hard, and I have become imbued with great respect for mathematics, the subtler part of which I had in my simple-mindedness regarded as pure luxury until now. Compared with this problem, the original relativity is child's play."*

Einstein had entered upon what were to be three years of the most intense work of his life. We have now to consider what he was up to.

---

[14] See *Subtle is the Lord*,[3] p. 210.

[15] Grossmann, you must help me or else I'll go crazy!

[16] Grossmann had not previously published in the areas in question, but had a good academic's familiarity with developments in his field. Einstein, on the other hand, did not possess a deep command of the literature, and was unaware that aspects of his train of thought had been anticipated decades before. In 1854 Riemann (1826–1866), in a *Habilitation* lecture entitled "Über die Hypothesen welche der Geometrie zugrunde liegen," had suggested that matter might be the cause of geometrical structure, and had in support of that conception described the outlines of "Riemannian geometry." That work was not published until 1867—the year following Riemann's death. In 1873 Clifford (1845–1879) arranged for an English translation of Riemann's essay to be published in *Nature*, and in his own "On the space-theory of matter" (1876) carried the idea even further: by the time he wrote Chapter 4 of *The Common Sense of the Exact Sciences* (1885) he was prepared to argue that not only mechanics but also electrodynamics—the whole of classical physics—are manifestations of the curvature of space. Similar ideas (if somewhat differently motivated) were advanced by Hertz (1857–1894) in his *The Principles of Mechanics* (1894). But these prescient thinkers worked in ignorance of special relativity, so contemplated the physical geometry not of spacetime but of space. Nor were they in position to draw guidance from the Principle of Equivalence.

General relativity lies on the other side of the mathematical thicket before which we now stand, and through which we—like Einstein in 1912, and like every student of gravitation since 1915—are obliged now to thread our way. The path was not yet clearly marked in Einstein's day (though the thicket had been in place—neglected by the generality of mathematicians—since before he was born) but has by now been very clearly mapped by any number of authors. I hesitate to add to that vast literature. Were it my option I would say "Find a book, as highbrow or lowbrow, as abstractly elegant or specifically concrete as seems most comfortable to you, and come back when you have mastered it." But that might take a while, and when you did come back we would almost certainly find that we had developed a language/notation problem. So... I attempt now to visit the principal landmarks, and to describe them in terms calculated to serve my immediate practical needs but which make no claim to mathematical modernity.

MATHEMATICAL DIGRESSION

Tensor analysis on Riemannian manifolds

When Riemann devised "Riemannian geometry" (1854)—which he did at the instigation of Gauss, who had selected the least favored of the three *Habilitation* topics proposed by Riemann—he built upon earlier work done by Gauss himself, and was influenced also by the then fairly recent non-Euclidean geometries of Lobachevsky and Bolyai (∼1830)... but managed to say what he had to say entirely without reference to "tensor analysis" (which hadn't been invented yet). The first steps toward the creation of the latter subject were taken by Elwin Christoffel (1829–1900), who was a disciple of Riemann, and in 1882 was motivated to invent what we now call the "covariant derivative." Gregorio Ricci-Curbastro (1853–1925) was for the last forty-five years of his life a professor of mathematical physics at the University of Padua,[17] and it was there that, during the years 1884–1894 and drawing inspiration from Riemann and Christoffel, he single-handedly invented what he called the "absolute differential calculus." During the latter phases of that work he was joined by his student, Tullio Levi-Civita (1873–1941), and together they wrote the monograph "Méthodes de calcul différentiel absolute et leurs applications" (Mathematische Annalen 1900) which brought tensor analysis to a recognizably modern form (though it was not until 1917 that Levi-Civita invented the important concept of "parallel transport"). This accomplishment was not widely applauded, and in some quarters inspired hostility; it was in reaction to Ricci's work that in 1899 Élie Cartan published the paper which laid the foundation for what was to become the "exterior calculus." For an account of

[17] Early in his career he had, at the instigation of Betti, published a memoir in *Nuovo Cimento* which introduced Italian physicists to the electrodynamics of Maxwell, and during a post-doctoral year in Munich (1877/8) he had come under the influence of Felix Klein.

the relationship between tensor analysis and the exterior calculus—an account which contains still some echo of that ancient tension—see §1.2 in H. Flanders' *Differential Forms* (1963).

I offer the preceding thumbnail history in order to underscore this fact: in opting to construct a *simultaneous* account of the relevant essentials of Riemannian geometry and tensor analysis I am melding two semi-independent subjects, one of which is fully thirty years older than the other (but both of which had been in place for nearly twenty years by the time Einstein was motivated to draw upon them).

Metrically connected manifolds. To start with the concrete: let x^1, x^2, x^3 refer to a Cartesian fraame in Euclidean 3-space. To describe the distance between two neighboring points write

$$\begin{aligned}(ds)^2 &= (dx^1)^2 + (dx^2)^2 + (dx^3)^2 \\ &= \delta_{ij} dx^i dx^j \quad \text{with} \quad \|\delta_{ij}\| \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\end{aligned} \tag{9}$$

and agree to call δ_{ij} the "metric connection." Let equations $x^i = x^i(y^1, y^2, y^3)$ serve to describe the introduction $x \to y$ of an arbitrary (and in the general case curvilinear) recoordinatization of 3-space.[18] Evidently

$$\begin{aligned}x \to y \quad \text{induces} \quad dx^i \to dy^i = & M^i{}_j dx^j \\ & M^i{}_j \equiv \partial y^i / \partial x^j\end{aligned} \tag{10.1}$$

which is to say:

> *coordinate differentials transform as components of a*
> *<u>contravariant tensor</u> of first rank (contravariant vector)*

The inverse transformation $x \leftarrow y$ induces

$$\begin{aligned}dx^i = & W^i{}_j dy^j \leftarrow dy^i \\ & W^i{}_j \equiv \partial x^i / \partial y^j\end{aligned} \tag{10.2}$$

Consistently with the elementary observation that $x \to y \to x$ must be the identity transformation, we (by the chain rule) have

$$W^i{}_a M^a{}_j = \textstyle\sum_a \frac{\partial x^i}{\partial y^a} \frac{\partial y^a}{\partial x^j} = \partial x^i / \partial x^j = \delta^i{}_j$$

which is to say: $\mathbb{W}\,\mathbb{M} = \mathbb{I}$. Our assumption that $x \to y$ is invertible can be expressed $\det \mathbb{M} \neq 0$, which assures the existence of $\mathbb{W} = \mathbb{M}^{-1}$.

[18] We require $x \to y$ to be invertible (maybe not globally, but at least) in a neighborhood containing the point P where, for the purposes of this discussion, we have elected to live.

The geometrical *meaning* of $(ds)^2$ is clearly independent of coordinatized language we elect to speak when *describing* it. At (9) we spoke in x-language. To say the same thing in y-language we write

$$\begin{aligned}(ds)^2 =& g_{ij}(y)dy^i dy^j \\ & g_{ij}(y) \equiv \delta_{ab}\tfrac{\partial x^a}{\partial y^i}\tfrac{\partial x^b}{\partial y^j}\end{aligned} \tag{11}$$

which is to say:

$$x \to y \quad \text{induces} \quad \delta_{ij} \to g_{ij} = W^a{}_i W^b{}_j \delta_{ab}$$

In words,

the metric connection transforms as
a <u>covariant tensor</u> of second rank

In matrix notation we have $\|\delta_{ij}\| \to \|g_{ij}\| = \mathbb{W}^\intercal \|\delta_{ab}\| \mathbb{W}$ from which it follows that

$$\begin{aligned}x \to y \quad \text{induces} \quad \det\|\delta_{ij}\| \to \det\|g_{ij}\| =& W^2 \cdot \det\|\delta_{ab}\| \\ & W \equiv \det \mathbb{W}\end{aligned}$$

In words,

the determinant $g \equiv \|g_{ij}\|$ of the metric connection
transforms as a scalar <u>density</u> of weight $w = 2$

More generally, we would say of the multiply-indexed objects $X^{ijk}{}_{mn}$ that they transform "as components of a mixed tensor of

- contravariant rank 3 (number of superscripts)
- covariant rank 2 (number of subscripts) and
- weight w"

if and only they respond to $x \to y$ by the rule

$$X^{ijk}{}_{mn} \to Y^{ijk}{}_{mn} = W^w \cdot M^i{}_a M^j{}_b M^k{}_c W^d{}_m W^e{}_n X^{abc}{}_{de} \tag{12}$$

which generalizes straightforwardly to arbitrary covariant/contravariant rank. It remains to be established that geometry/physics present a vast number of objects which *do* transform by this rule (as well as a population of multiply-indexed objects made all the more interesting by the fact that they don't!). And it is important to appreciate the significance of the "as components of" in the sentence which led to (12); it is important, that is to say, not to confuse the geometrical/physical object $\mathfrak{X}$ with the set $X^{ijk}{}_{mn}$ of numbers which, *relative to a coordinate system*, serve to describe it (its "coordinates").

From (12) follow a number of propositions—collectively, the subject matter of "tensor algebra"—which are in each instance either self-evident or so easy to prove that I state them without proof:

• Tensors[19] can be added/subtracted (to yield again a tensor) if and only if they possess the same covariant/contravariant ranks and weight (for otherwise they would come unstuck when transformed).

• $A^{\cdots}{}_{\cdots} = B^{\cdots}{}_{\cdots}$ means $A^{\cdots}{}_{\cdots} - B^{\cdots}{}_{\cdots} = 0$, which if valid in one coordinate system is, by the design of (12), clearly valid in all. Such tensor equations require that $A^{\cdots}{}_{\cdots}$ and $B^{\cdots}{}_{\cdots}$ have the same ranks and weight, and provide coordinatized expression of statements of the form $\mathfrak{A} = \mathfrak{B}$.

• Tensors can be multiplied (to yield again a tensor) even if they possess distinct ranks and weights; the resulting tensor will have
  * contravariant/covariant rank equal to the sum of the respective ranks of the factors
  * weight equal to the sum of the weights of the factors.

• If $A^{\cdots i\cdots}{}_{\cdots j\cdots}$ is a tensor density of contravariant rank r and covariant rank s then $A^{\cdots a\cdots}{}_{\cdots a\cdots}$ transforms as a tensor density of the same weight, and of the respective ranks $r-1$ and $s-1$. We say the ^{i} has been "contracted" into the $_j$. Attempted contraction of a superscript into a superscript (or of a subscript into subscript) would, on the other hand, yield an object which fails to transform tensorially.

• It makes transform-theoretic good sense to say of a tensor that it is symmetric/antisymmetric with respect to some specified pair of superscripts/subscripts

$$A^{\cdots i\cdots j\cdots}{}_{\cdots} = \pm A^{\cdots j\cdots i\cdots}{}_{\cdots} \quad \text{or} \quad A^{\cdots}{}_{\cdots i\cdots j\cdots} = \pm A^{\cdots}{}_{\cdots j\cdots i\cdots}$$

but statements of the form $A^{\cdots i\cdots}{}_{\cdots j\cdots} = \pm A^{\cdots j\cdots}{}_{\cdots i\cdots}$ come unstuck when transformed.

• It is in this light remarkable that Kronecker's mixed tensor—defined

$$\delta^i{}_j \equiv \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

in some coordinate system—retains that description in all coordinate systems.

Returning again to (11), we are in position now to understand the coordinate-independence of $(ds)^2$ to be a result of our having contracted the second rank product $dy^i dy^j$ of a pair of contravariant vectors into a second rank covariant tensor g_{ij}. And to observe that (as a result of the "Pythagorean symmetry" originally attributed to δ_{ij}) the metric tensor $g_{ij}(y)$ will in all coordinate systems be symmetric:

$$g_{ij} = g_{ji}$$

[19] ...of, it need hardly be added, the *same dimension*. On says of a tensor that it is "N-dimensional" if the indices range on $\{1, 2, \ldots, N\}$.

To $g_{ij}(y)$ we assign a contravariant companion $g^{ij}(y)$, defined by the square array of linear equations

$$g^{ia}g_{aj} = \delta^i{}_j$$

The numbers g^{ij} are in effect just the elements of the matrix $\|g_{ij}\|^{-1}$. We observe that necessarily

$$g^{ij} = g^{ji}$$

and that $\det\|g^{ij}\| = 1/g$ transforms as a scalar density of weight $w = -2$.

Recalling that at the beginning of this conversation we placed ourselves in Euclidean 3-space, let the y-coordinates be, for purposes of illustration, spherical: then

$$\left.\begin{aligned} x^1 &= r\,\sin\theta^1\cos\theta^2 \\ x^2 &= r\,\sin\theta^1\sin\theta^2 \\ x^3 &= r\cos\theta^1 \end{aligned}\right\} \tag{13}$$

give

$$\begin{aligned} dx^1 &= \sin\theta^1\cos\theta^2\cdot dr + r\cos\theta^1\cos\theta^2\cdot d\theta^1 - r\sin\theta^1\sin\theta^2\cdot d\theta^2 \\ dx^2 &= \sin\theta^1\sin\theta^2\cdot dr + r\cos\theta^1\sin\theta^2\cdot d\theta^1 + r\sin\theta^1\cos\theta^2\cdot d\theta^2 \\ dx^3 &= \qquad\cos\theta^1\cdot dr - \qquad r\sin\theta^1\cdot d\theta^1 \end{aligned}$$

whence (squaring and simplifying; it would in the present context be cheating to simply read from a figure)

$$(ds)^2 = (dr)^2 + r^2(d\theta^1)^2 + (r\sin\theta^1)^2(d\theta^2)^2$$

Let us now take up residence *on the 2-dimensional surface* of the sphere of radius R (see the following figure); we then have

$$\begin{aligned} (ds)^2 &= R^2(d\theta^1)^2 + (R\sin\theta^1)^2(d\theta^2)^2 \\ &= g_{ij}(\theta)d\theta^i d\theta^j \quad \text{with} \quad \|g_{ij}\| = \begin{pmatrix} R^2 & 0 \\ 0 & (R\sin\theta^1)^2 \end{pmatrix} \end{aligned} \tag{14}$$

which describes in θ-coordinates the non-Euclidean metric structure that the spherical surface has inherited from the enveloping Euclidean 3-space. Notice that

$$g \equiv \det\|g_{ij}\| = R^4\sin^2\theta^1 \text{ vanishes at the poles}$$

where g^{ij} therefore fails to exist, essentially because (13) becomes non-invertible on the polar axis.

We have come upon a particular instance of a class of objects—surfaces Σ^2 embedded in Euclidean 3-space E^3—which (especially with regard to their curvature properties) were the subject of Gauß' pioneering *Disquisitiones generales circa superficies curvas* (1827). It is to expose the magnitude of Riemann's accomplishment that I look now very briefly to Gauß' "theory of surfaces."

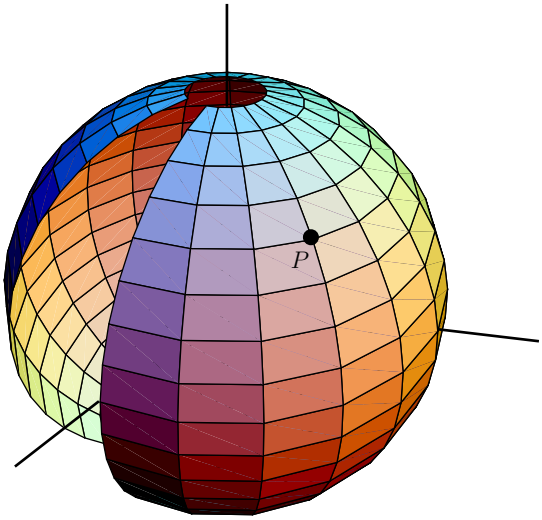


FIGURE 3: *Spherical surface from which polar caps and a sector have been excised to avoid complications which result from the circumstance that the* $\{\theta^1, \theta^2\}$ *coordinate system (13) becomes singular at the poles (where* $\cos\theta^1 = \pm 1$*) and is periodic in* θ^2*. The surface inherits its metric properties from the enveloping Euclidean space.*

Write $\boldsymbol{x}(u, v)$ to present a parametric description of such a surface Σ^2 (which u and v serve to coordinatize). Let P and Q mark a pair of neighboring points on Σ^2. To describe the squared Euclidean length of the interval separating Q from P, Gauss writes

$$(ds)^2 = d\boldsymbol{x}\cdot d\boldsymbol{x} \quad \text{with} \quad d\boldsymbol{x} = \frac{\partial \boldsymbol{x}}{\partial u}\, du + \frac{\partial \boldsymbol{x}}{\partial v}\, dv + \cdots \equiv \boldsymbol{x}_u\, du + \boldsymbol{x}_v\, dv$$

and obtains

$$\begin{aligned} (ds)^2 &= \boldsymbol{x}_u\cdot\boldsymbol{x}_u\, dudu + 2\boldsymbol{x}_u\cdot\boldsymbol{x}_v\, dudv + \boldsymbol{x}_v\cdot\boldsymbol{x}_v\, dvdv \\ &\equiv E\, dudu + 2F\, dudv + G\, dvdv \\ &\equiv \text{FIRST FUNDAMENTAL FORM} \end{aligned} \tag{15.1}$$

of which (14) presents a particular instance. Carrying the expansion of $d\boldsymbol{x}$ to higher order

$$d\boldsymbol{x} = \big\{\boldsymbol{x}_u\, du + \boldsymbol{x}_v\, dv\big\} + \tfrac{1}{2!}\big\{\boldsymbol{x}_{uu}\, dudu + 2\boldsymbol{x}_{uv}\, dudv + \boldsymbol{x}_{vv}\, dvdv\big\} + \cdots$$

Gauss observes that the leading term on the right lies in the plane tangent to Σ^2 at P, and that the second term provides lowest-order indication of how Σ^2

curves away from that plane; taking $\hat{\boldsymbol{n}}$ to be the unit normal to the tangent plane, Gauss constructs

$$\begin{aligned}\hat{\boldsymbol{n}}\cdot d\boldsymbol{x} &= \tfrac{1}{2}\big\{\hat{\boldsymbol{n}}\cdot\boldsymbol{x}_{uu}\,dudu + 2\hat{\boldsymbol{n}}\cdot\boldsymbol{x}_{uv}\,dudv + \hat{\boldsymbol{n}}\cdot\boldsymbol{x}_{vv}\,dvdv\big\}\\ &\equiv \tfrac{1}{2}\big\{e\,dudu + 2f\,dudv + g\,dvdv\big\}\\ &\equiv \tfrac{1}{2}\cdot\text{SECOND FUNDAMENTAL FORM}\end{aligned} \tag{15.2}$$

where $\hat{\boldsymbol{n}}$ can be described

$$\begin{aligned}\hat{\boldsymbol{n}} = \frac{\boldsymbol{x}_u\times\boldsymbol{x}_v}{\sqrt{(\boldsymbol{x}_u\times\boldsymbol{x}_v)\cdot(\boldsymbol{x}_u\times\boldsymbol{x}_v)}} &= \boldsymbol{x}_u\times\boldsymbol{x}_v\Big/\sqrt{\det\begin{pmatrix}\boldsymbol{x}_u\cdot\boldsymbol{x}_u & \boldsymbol{x}_u\cdot\boldsymbol{x}_v\\ \boldsymbol{x}_v\cdot\boldsymbol{x}_u & \boldsymbol{x}_v\cdot\boldsymbol{x}_v\end{pmatrix}}\\ &= \boldsymbol{x}_u\times\boldsymbol{x}_v\Big/\sqrt{EG-F^2}\end{aligned}$$

Possession of the first and second fundamental forms placed Gauss in position to construct an account of the *curvature of* Σ^2 *at* P which builds upon the elementary theory of the curvature of plane curves. Consider the planes which stand normal to Σ^2 at P (in the sense that each contains $\hat{\boldsymbol{n}}$). The intersection of Σ^2 with such a plane inscribes a curve upon that plane, which at P has curvature κ. Gauss shows[20] that the least/greatest values of κ to be found within that set of planes can be obtained as the respective roots of a quadratic polynomial

$$(EG-F^2)\kappa^2 - (Eg-2Ff+Ge)\kappa + (eg-f^2) = 0$$

which combines information written into the fundamental forms. Let those roots be denoted κ_1 and κ_2; one has then the definitions

$$\begin{aligned}\text{GAUSSIAN CURVATURE} \quad &: \quad K \equiv \kappa_1\kappa_2 = \frac{eg-f^2}{EG-F^2}\\ \text{MEAN CURVATURE} \quad &: \quad M \equiv \tfrac{1}{2}(\kappa_1+\kappa_2) = \frac{Eg-2Ff+ge}{2(EG-F^2)}\end{aligned} \tag{16}$$

which serve in their respective ways to describe the curvature of Σ^2 at P. Looking to the dimensional implications of the definitions (15) we find that $[\kappa] = (\text{length})^{-1}$, as required by the familiar relation

$$\text{curvature} = \frac{1}{\text{radius of curvature}}$$

[20] See, for example, §§32–38 in H. Lass, *Vector & Tensor Analysis* (1950) or §§9.14–9.16 in K. Rektorys, *Survey of Applicable Mathematics* (1969) for the elementary details, and accessible surveys of related matters. More modern discussions—such, for example, as can be found in John McCleary, *Geometry from a Differentiable Viewpoint* (1994), Chapter 9 or R. Darling, *Differential Forms & Connections* (1994) §§4.7–4.11—tend generally to be vastly more elegant, but (in proportion to their self-conscious modernism) to impose such heavy formal demands upon the reader as to be almost useless except to people who have already gained a geometrical sense of the subject from other sources.

Look back again, by way of illustration, to the spherical surface shown in Figure 3. We have

$$\boldsymbol{x}(u,v) = R\,\hat{\boldsymbol{n}}(u,v) \quad \text{with} \quad \hat{\boldsymbol{n}} = \begin{pmatrix} \sin u \cos v \\ \sin u \sin v \\ \cos u \end{pmatrix}$$

giving

$$\boldsymbol{x}_u = R \begin{pmatrix} +\cos u \cos v \\ +\cos u \sin v \\ -\sin u \end{pmatrix}, \qquad \boldsymbol{x}_v = R \begin{pmatrix} -\sin u \sin v \\ +\sin u \cos v \\ 0 \end{pmatrix}$$

$$\boldsymbol{x}_{uu} = -\boldsymbol{x}, \qquad \boldsymbol{x}_{uv} = R \begin{pmatrix} -\cos u \sin v \\ +\cos u \cos v \\ 0 \end{pmatrix}, \qquad \boldsymbol{x}_{vv} = R \begin{pmatrix} -\sin u \cos v \\ -\sin u \sin v \\ 0 \end{pmatrix}$$

from which we compute

$$\begin{aligned} E &= R^2 & e &= -R \\ F &= 0 & f &= 0 \\ G &= R^2 \sin^2 u & g &= -R \sin^2 u \end{aligned}$$

from which it follows by (16) that

$$K = \frac{1}{R^2} \quad : \quad \text{all values of } u \text{ and } v$$

The specialness of the result reflects, in an obvious sense, the specialness of spheres.

The functions

$$\begin{aligned} g_{uu}(u,v) &= E(u,v) \\ g_{uv}(u,v) = g_{vu}(u,v) &= F(u,v) \\ g_{vv}(u,v) &= G(u,v) \end{aligned}$$

are "intrinsic to Σ^2" in the sense that they are accessible to determination by a flat mathematician who has a good understanding of the Euclidean geometry of E^2 but no perception of the enveloping E^3. The $\hat{\boldsymbol{n}}$ which enters into the definitions of $e(u,v)$, $f(u,v)$ and $g(u,v)$ presents, on the other hand, an explicit reference to the enveloping space. It is, in this light, remarkable—in Latin: *egregium*—that

Gaussian curvature is an *intrinsic* property of Σ^2

which is the upshot of Gauss' THEOREMA EGREGIUM, the capstone of his *Disquisitiones*. The technical point is that e, f and g can be expressed in terms of E, F and G; when this is done, (16) becomes

$$K = -\frac{1}{4D^4}\begin{vmatrix} E & E_u & E_v \\ F & F_u & F_v \\ G & G_u & G_v \end{vmatrix} - \frac{1}{2D}\left\{\frac{\partial}{\partial v}\frac{E_v - F_u}{D} - \frac{\partial}{\partial u}\frac{F_v - G_u}{D}\right\} \qquad (17)$$

with $D \equiv \sqrt{EG - F^2}$.[21] Gauss stressed also the fact that (17) is structurally invariant with respect to recoordinatizations

$$\left.\begin{matrix} u \\ v \end{matrix}\right\} \longrightarrow \left\{\begin{matrix} u' = u'(u,v) \\ v' = v'(u,v) \end{matrix}\right.$$

of the surface Σ^2.

To consult the literature is to encounter the names of people like G. Mainardi (1800–1879), D. Codazzi (1824–1875) and O. Bonnet (1819–1892) who were productive participants in the flurry of geometrical activity which followed publication of *Disquisitiones*. But it is clear in retrospect—and was clear already to Gauss[22]—that the most profoundly creative of that second generation of geometers was Riemann. Gauss had been led to geometry from physical geodesy, which may account for why he concentrated on the geometry of surfaces in 3-space (as did those who followed in his steps). Riemann, on the other hand, imagined himself to be exploring the "foundations of geometry," and in doing so to be pursuing a complex physico-philosophical agenda;[23] he found it natural[24] to look upon Gaussian surfaces as special cases of much more general (N-dimensional) structures. "Manifolds" he called them, the properties of which were internal to themselves—developed without reference to any enveloping Euclidean space.

[21] In the spherical case we would be led on this basis to write

$$K = -0 - \frac{1}{2R^2 \sin u}\left\{\frac{\partial}{\partial v}0 - \frac{\partial}{\partial u}(-2\cos u)\right\} = \frac{1}{R^2}$$

which checks out.

[22] Gauss died in 1855, only one year after Riemann—at age 28—had (after delays caused by Gauss' declining health) presented the geometrical lecture in which Gauss found so much to praise. By 1860 Riemann himself was dead.

[23] See E. T. Bell captures the flavor of that agenda in Chapter 26 of his *Men of Mathematics* (1937). Or see Michael Spivak's translation of "Riemann's *Habilatationsvortrag*: On the hypotheses which lie at the foundations of geometry," which is reprinted in McCleary,[20] which is pretty heavy sledding, but ends with the remark that "This leads us away into the domain of another science, the realm of physics, into which the nature of the present occasion does not allow us to enter."

[24] There is no accounting for genius, but could he have been influeced by "Riemannian surfaces of N-sheets" which had played a role in his dissertation, "Grundlagen für eine allegemeine Theorie der Functionen einer veränderlichen complexen Grösse" (1851)?

With a vagueness worthy of Riemann himself, I will understand an N-dimensional manifold to be an "$(x^1, x^2, \ldots, x^N)$-coordinatized continuum" which is sufficiently structured to permit us to do the things we want to do. The manifold becomes a *Riemannian manifold* when endowed with functions

$$g_{ij}(x^1, x^2, \ldots, x^N) \quad : \quad i, j \in \{1, 2, \ldots, N\},\ g_{ij} = g_{ji}$$

which permit one to write

$$(ds)^2 = g_{ij}(x)\, dx^i dx^j \tag{18}$$

to lend postulated *metric structure* to the manifold. In (18) we maintain the form of (12) but abandon the notion that (18) is the curvilinear expression of some more primitive metric axiom (Euclidean metric, either of the manifold itself or—as it was at (14), and always was for Gauss—of a Euclidean space within which the manifold is imagined to be embedded). Riemann retains the service of the "first fundamental form," but his program obligates him to abandon the "second fundamental form," and therefore to find some way to circumvent the mathematics which for Gauss culminated in the Theorema Egregium.

Somewhat idiosyncratically, I use the word "connection" to refer to *any* ancillary device we "smear on a manifold" in order to permit us to *do* things there, and say of a manifold $\mathcal{M}$ that has been endowed with a $g_{ij}(x)$ that it has been "metrically connected."[25]

The coordinate-independence of (18) requires that g_{ij} transform as a tensor. At each of the points P of $\mathcal{M}$ we erect, as our mathematical/physical interest may dictate, also populations of other tensors of various ranks and weights. Those live in multivector spaces which are "tangent" to $\mathcal{M}$ at P. If X^i refers (in x-coordinates) to a contravariant vector defined at P, then $g_{ia}X^a$ and $g_{ai}X^a$ refer similarly to covariant vectors defined at P. And these are, by the symmetry of g_{ij}, the *same* covariant vector, which we may agree to denote X_i. Moreover, $g^{ia}X_a = g^{ia}g_{ab}X^b = \delta^i{}_b X^b = X^i$ gives back the contravariant vector from which we started. The idea extends naturally to tensors of arbitrary rank and weight

$$X^{\cdots}{}_{i}{}^{\cdots}{}_{\cdots} = g_{ia}X^{\cdots a\cdots}{}_{\cdots} \quad ; \quad X^{\cdots i\cdots}{}_{\cdots} = g^{ia}X^{\cdots}{}_{a}{}^{\cdots}{}_{\cdots}$$

On metrically connected manifolds we can agree to press the metric connection g_{ij} into secondary service as the *universal index manipulator*. The fundamental Reimannian axiom (18) can by this convention be written in a way

$$(ds)^2 = dx_i\, dx^i$$

which renders g_{ij} itself covert.

[25] My usage is arguably consistent with that employed by Schrödinger (See Chapter 9 in his elegant little book, *Spacetime Structure*(1954)), but departs from that favored by most differential geometers. As used by the latter, the concept originates in work (1916) of Gerhard Hassenberg, a German set theorist.

Geodesics. Let $x^i(t)$ describe a t-parameterized curve $\mathcal{C}$ which has been inscribed on $\mathcal{M}$. Assume

$$x^i(0) = \text{coordinates of a point } P$$
$$x^i(1) = \text{coordinates of a point } Q$$

i.e., that the curve $\mathcal{C}_{P\to Q}$ links P to Q. Riemann's axiom (18) places us in position to write

$$\text{length of } \mathcal{C}_{P\to Q} = \int_0^1 \sqrt{g_{ab}v^a v^b}\, dt$$
$$v^a \equiv \tfrac{d}{dt}x^a(t)$$

"Geodesics" are curves of extremal length, and by straightforward appeal to the calculus of variations are found to satisfy

$$\left\{\frac{d}{dt}\frac{\partial}{\partial v^i} - \frac{\partial}{\partial x^i}\right\}\sqrt{g_{ab}v^a v^b} = 0 \tag{19.1}$$

Working out the implications of (19.1) is a task made complicated by the presence of the radical. Those complications[26] can, however, be circumvented; one can show without difficulty[27] that if $\dot{s} = 0$ —i.e., if

$$t = (\text{constant}) \cdot (\text{arc length}) + (\text{constant})$$

—then the $\sqrt{}$ can be discarded. So we adopt *arc-length parameterization* and obtain

$$\left\{\frac{d}{ds}\frac{\partial}{\partial u^i} - \frac{\partial}{\partial x^i}\right\}g_{ab}u^a u^b = 0 \tag{19.2}$$
$$u^a \equiv \tfrac{d}{ds}x^a(s)$$

Quick calculation gives $g_{ia}\frac{d}{ds}u^a + \frac{1}{2}\{\partial_a g_{ib} + \partial_b g_{ia} - \partial_i g_{ab}\}u^a u^b = 0$ which can be written

$$\tfrac{d}{ds}u^i + \left\{\begin{smallmatrix} i \\ ab \end{smallmatrix}\right\} u^a u^b = 0 \tag{20.2}$$
$$\left\{\begin{smallmatrix} i \\ ab \end{smallmatrix}\right\} \equiv g^{ij}\cdot\tfrac{1}{2}\{\partial_a g_{jb} + \partial_b g_{ja} - \partial_j g_{ab}\} \tag{21}$$

To recover the result to which we would have been led had we proceeded from (19.1)—i.e., had we elected to use *arbitrary* parameterization—we use $\frac{d}{ds} = (\dot{s})^{-1}\frac{d}{dt}$ (whence $u^i = (\dot{s})^{-1}v^i$) and obtain

$$\tfrac{d}{dt}v^i + \left\{\begin{smallmatrix} i \\ ab \end{smallmatrix}\right\} v^a v^b = v^i \tfrac{d}{dt}\log\tfrac{ds}{dt} \tag{20.1}$$
$$\tfrac{ds}{dt} = \sqrt{g_{ab}v^a v^b}$$

[26] See CLASSICAL DYNAMICS (1964), Chapter 2, p. 105.

[27] The simple argument can be found in "Geometrical mechanics: Remarks commemorative of Heinrich Hertz" (1994) at p. 14.

Comparison of (19.2) with 19.1 , (20.2) with (20.1) underscores the marked *simplification which typically results from arc-length parameterization.*

Covariant differentiation on affinely connected manifolds. If $X(x)$ responds to $x \to y$ as a scalar density of zero weight

$$X(x) \to Y(y) = X(x(y))$$

then

$$\frac{\partial Y}{\partial y^i} = \frac{\partial x^a}{\partial y^i}\frac{\partial X}{\partial x^a} \quad \text{which we abbreviate} \quad Y_{,i} = W^a{}_i X_{,a}$$

In short: the gradient of a scalar transforms tensorially (i.e., as a covariant vector, and is standardly held up as the exemplar of such an object). But if X_i transforms as a covariant vector ($Y_i = W^a{}_i X_a$) then

$$Y_{i,j} = W^a{}_i W^b{}_j X_{a,b} + X_a \cdot \frac{\partial^2 x^a}{\partial y^i \partial y^j}$$

shows that $X_{i,j}$ *fails to transform tensorially* (unless we impose upon $x \to y$ the restrictive requirement that $\partial^2 x^a/\partial y^i \partial y^j = 0$). A similar remark pertains generally: if $Y^{i_1 \cdots i_r}{}_{j_1 \cdots j_s} = W^w \cdot M^{i_1}{}_{a_1} \cdots M^{i_r}{}_{a_r} W^{b_1}{}_{j_1} \cdots W^{b_s}{}_{j_s} X^{a_1 \cdots a_r}{}_{b_1 \cdots b_s}$ then

$$Y^{i_1 \cdots i_r}{}_{j_1 \cdots j_s,k} = \big[W^w \cdot M^{i_1}{}_{a_1} \cdots M^{i_r}{}_{a_r} W^{b_1}{}_{j_1} \cdots W^{b_s}{}_{j_s}\big] W^c{}_k X^{a_1 \cdots a_r}{}_{b_1 \cdots b_s,c} + X^{a_1 \cdots a_r}{}_{b_1 \cdots b_s} \frac{\partial}{\partial y^k}\big[\text{etc.}\big]$$

This circumstance severely constrains our ability to do ordinary differential calculus on manifolds; it limits us on transformation-theoretic grounds to such "accidentally tensorial" constructs as $X_{i,j} - X_{j,i}$ in which the unwanted terms (by contrivance) cancel.[28]

There exists, however, an elegant work-around, devised early in the present century by Ricci and Levi-Civita (who harvested the fruit of a seed planted by Christoffel in 1869) and which has much in common with the spirit of gauge field theory. In place of ∂_j we study a *modified* operator $\mathcal{D}_j$, the action of which can, in the simplest instance, be described

$$\mathcal{D}_j X_i \equiv X_{i,j} - X_a \Gamma^a{}_{ij} \quad ; \quad \text{denoted } X_{i;j} \qquad (22)$$

($X_{i;j}$: semi-colon instead of comma)

The idea is to impose upon the "affine connection" $\Gamma^a{}_{ij}$ such transformation properties as are sufficient to *insure* that $X_{i;j}$ transform tensorially

$$\bar{X}_{i;j} = W^a{}_i W^b{}_j X_{a;b}$$

[28] For lists of such "accidentally tensorial constructs"—of which the exterior calculus provides a systematic account (and which are in themselves rich enough to support much of physics)—see §4 of "Electrodynamical Applications of the Exterior Calculus" (1996) and pp. 22-24 of Schrödinger.[25]

(It has at this point become more natural to write $x \to \bar{x}$ where formerly we wrote $x \to y$.) This by

$$\begin{aligned}\bar{X}_{i,j} - \bar{X}_a \bar{\Gamma}^a{}_{ij} &= \Big\{ W^a{}_i W^b{}_j X_{a,b} + X_a \cdot \frac{\partial^2 x^a}{\partial y^i \partial y^j} \Big\} - W^b{}_a X_b \bar{\Gamma}^a{}_{ij} \\ &= W^a{}_i W^b{}_j X_{a,b} - X_c \Big\{ W^c{}_a \bar{\Gamma}^a{}_{ij} - \frac{\partial^2 x^c}{\partial y^i \partial y^j} \Big\} \\ &= W^a{}_i W^b{}_j \big(X_{a,b} - X_c \Gamma^c{}_{ab} \big)\end{aligned}$$

requires that $W^c{}_a \bar{\Gamma}^a{}_{ij} - \partial^2 x^c / \partial y^i \partial y^j = W^a{}_i W^b{}_j \Gamma^c{}_{ab}$. Multiplication by $M^k{}_c$ leads to the conclusion that $\Gamma^k{}_{ij}$ must transform

$$\Gamma^k{}_{ij} \to \bar{\Gamma}^k{}_{ij} = M^k{}_c W^a{}_i W^b{}_j \Gamma^c{}_{ab} + \frac{\partial y^k}{\partial x^c} \frac{\partial^2 x^c}{\partial y^i \partial y^j} \tag{23}$$

if it is to do the job we require of it. To assign natural meaning to $X^i{}_{;j}$ we shall REQUIRE that *for weightless scalars "covariant differentiation" reduces to ordinary differentiation*

$$X_{;i} = X_{,i} \tag{24.1}$$

and that for *for tensor products covariant differentiation satisfies the product rule*

$$(X^{\cdots}{}_{\cdots} Y^{\cdots}{}_{\cdots})_{;i} = X^{\cdots}{}_{\cdots;i} Y^{\cdots}{}_{\cdots} + X^{\cdots}{}_{\cdots} Y^{\cdots}{}_{\cdots;i} \tag{24.2}$$

Look in this light to the contracted product $Y_a X^a$; we have

$$(Y_{a,j} - Y_i \Gamma^i{}_{aj}) X^a + Y_i X^i{}_{;j} = (Y_i X^i)_{,j} = Y_{a,j} X^a + Y_i X^i{}_{,j}$$

for all Y_i, from which we obtain the second of the following equations (the first being simply a repeat of (22)):

$$\left.\begin{aligned} X_{i;j} &= X_{i,j} - X_a \Gamma^a{}_{ij} \\ X^i{}_{;j} &= X^i{}_{,j} + X^a \Gamma^i{}_{aj} \end{aligned}\right\} \tag{25.1}$$

A somewhat more intricate argument—which I omit[29]—leads to the conclusion that for scaral *densities* the natural thing to write is

$$X_{;i} = X_{,i} - wX\Gamma^a{}_{ai} \tag{25.2}$$

which gives back (24.1) in the case $w = 0$. Using (24) and (25) in combination one can describe the covariant derivatives of tensors of all ranks and weights; for example, we find

$$X^{ij}{}_{k;l} = X^{ij}{}_{k,l} + X^{aj}{}_k \Gamma^i{}_{al} + X^{ia}{}_k \Gamma^j{}_{al} - X^{ij}{}_a \Gamma^a{}_{kl} - wX^{ij}{}_k \Gamma^a{}_{al}$$

[29] See Chapter 2, p. 59 in the notes recently cited,[26] or Schrödinger,[25] p. 32.

which—by construction—transforms as a mixed tensor density of

- unchanged contravariant rank,
- augmented covariant rank, and
- unchanged weight.

All of which becomes available as a mechanism for obtaining tensors by the (covariant) differentiation of tensors, and for constructing transformationally well-behaved tensorial differential equations…if and only if the underlying manifold $\mathcal{M}$ has been endowed with an affine connection; i.e., if and only if (in some coordinate system) functions ${\Gamma^k}_{ij}(x)$ have been prescribed at every point, in which case we say that $\mathcal{M}$ is "affinely connected." Differentiation of objects attached to such a manifold is *relative to the prescribed affine connection*: replace one affine connection with another, and the meaning of all derivatives changes. Description of ${\Gamma^k}_{ij}(x)$ requires the specification of N^3 functions; its description in other coordinate systems (reached by $x \to \bar{x}$) is then accomplished by appeal to (23), in connection with which we observe that

- ${\Gamma^k}_{ij}(x) \to {\bar{\Gamma}^k}_{ij}(\bar{x})$ is tensorial only if $x \to \bar{x}$ is linear (as, we note in passing, are the inhomogeneous Lorentz transformations); in more general cases ${\Gamma^k}_{ij}(x)$ transforms distinctively—"like an affine connection."

- The rule (23) is "transitive" in the sense that if ${\Gamma^k}_{ij}(x) \to {\bar{\Gamma}^k}_{ij}(\bar{x})$ and ${\bar{\Gamma}^k}_{ij}(\bar{x}) \to {\bar{\bar{\Gamma}}^k}_{ij}(\bar{\bar{x}})$ conform to it, then so does ${\Gamma^k}_{ij}(x) \to {\bar{\bar{\Gamma}}^k}_{ij}(\bar{\bar{x}})$

- If ${\Gamma^k}_{ij}(x)$ vanishes in some coordinate system (call it the x-system) then it is given in other coordinate systems by

$$ {\bar{\Gamma}^k}_{ij}(\bar{x}) = \frac{\partial \bar{x}^k}{\partial x^c} \frac{\partial^2 x^c}{\partial \bar{x}^i \partial \bar{x}^j} $$

  which is ij-symmetric.

- Let ${\Gamma^k}_{ij}$ be resolved ${\Gamma^k}_{ij} = \frac{1}{2}({\Gamma^k}_{ij} + {\Gamma^k}_{ji}) + \frac{1}{2}({\Gamma^k}_{ij} - {\Gamma^k}_{ji})$ into its ij-symmetric and ij-antisymmetric (or "torsional") parts. It follows from (23) that, while the symmetric part transforms "like an affine connection," the torsional part transforms *tensorially*.[30]

[30] In Riemannian geometry, and in gravitational theories based upon it, one (tacitly) assumes the affine connection to be "torsion free" (i.e., that ${\Gamma^k}_{ij}$ is symmetric), and *MTW* remark (p. 250) that to assume otherwise would be inconsistent with the Principle of Equivalence. But, beginning in the 1920's, Einstein and others studied various generalizations of general relativity which involved relaxation of that assumption. For a useful summary of that work see §§17d&e in Pais.[3] According to MTW (§39.2), all such generalizations can now be dismissed on observational grounds except (possibly) for the torsional theory described by E. Cartan in 1922/23.

On metrically connected manifolds there exists a "natural" (symmetric) affine connection. It arises when one imposes the requirement that

- covariant differentiation and
- index manipulation

be *compatible* operations, performable in either order:

$$(g_{ia}X^{\cdots a\cdots}{}_{\cdots})_{;k} = g_{ia}(X^{\cdots a\cdots}{}_{\cdots;k})$$

This, by (24.2), amounts to requiring that $g_{ij;k} = 0$; i.e., that

$$g_{ij,k} - g_{aj}\Gamma^a{}_{ik} - g_{ia}\Gamma^a{}_{jk} = 0$$

Let this equation be notated $\Gamma_{jik} + \Gamma_{ijk} = g_{ij,k}$ and draw upon the symmetry assumption to write Γ_{jki} in place of Γ_{jik}. Cyclic permutation on ijk leads then to a trio of equations which can be displayed

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \Gamma_{ijk} \\ \Gamma_{jki} \\ \Gamma_{kij} \end{pmatrix} = \begin{pmatrix} g_{jk,i} \\ g_{ki,j} \\ g_{ij,k} \end{pmatrix}$$

from which it follows by matrix inversion that

$$\begin{pmatrix} \Gamma_{ijk} \\ \Gamma_{jki} \\ \Gamma_{kij} \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} g_{jk,i} \\ g_{ki,j} \\ g_{ij,k} \end{pmatrix}$$

Making free use of metric symmetry ($g_{ij} = g_{ji}$) we are led thus—by an argument which is seen to hinge on the "permutation trick" employed already once before (see again the derivation of (2–83))—to cyclic permutations of the following basic statement:

$$\Gamma_{kij} = g_{ka}\Gamma^a{}_{ij} = \tfrac{1}{2}\big\{g_{ki,j} + g_{kj,i} - g_{ij,k}\big\} \tag{26.1}$$

$$\Gamma^k{}_{ij} = g^{ka} \cdot \tfrac{1}{2}\big\{g_{ai,j} + g_{aj,i} - g_{ij,a}\big\} \tag{26.2}$$

To compute the covariant derivatives of densities it helps to know also that

$$\Gamma^a{}_{ai} = \frac{\partial}{\partial x^i} \log \sqrt{g} \tag{26.3}$$

but the demonstration is somewhat intricate and will be postponed.[31]

The expressions which stand on the right side of (25) were introduced into the literature by Christoffel (1869); they are called "Christoffel symbols," and are (or used to be) standardly notated

$$[ij,k] \equiv \text{right side of (25.1)} \quad : \quad \text{Christoffel symbol of } 1^{\text{st}} \text{ kind}$$

$$\left\{ {k \atop ij} \right\} \equiv \text{right side of (25.2)} \quad : \quad \text{Christoffel symbol of } 2^{\text{nd}} \text{ kind}$$

[31] See p. 73 in some notes previously cited,[28] or (58) below.

The Christoffel symbol $\left\{ {k \atop ij} \right\}$ is the symmetric affine connection—the "Christoffel connection"—most natural to metrically connected manifolds (Riemannian manifolds), and its valuation is, according to (26), latent in specification of the metric $g_{ij}(x)$. It sprang spontaneously to our attention already at (20), when we were looking to the description of Riemannian geodesics. We confront therefore the question: What has covariant differentiation to do with geodesic design?

Parallel transport. When asked to "differentiate a tensor"—let it, to render the discussion concrete, be a covariant vector—one's first instinct might be to write

$$\frac{X_i(x+\delta x) - X_i(x)}{\delta x}$$

Such a program is, however, foredoomed. For, while it makes transformational good sense to write $Y_i(x) - X_i(x)$, it is not permissible to add/subtract vectors which are *attached to distinct points on the manifold* (vectors which therefore live in distinct vector spaces, and transform a bit differently from one another).

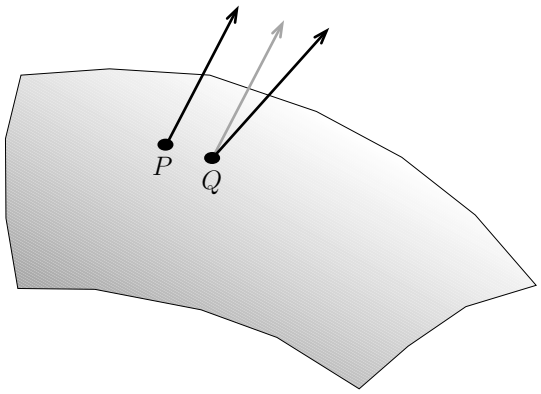


FIGURE 4: *The vector $X_i(x)$ is attached to the manifold at P, and $X_i(x+\delta x)$ at the neighboring point Q. The affine connection $\Gamma^k{}_{ij}(x)$ defines the sense in which the grey vector $\mathfrak{X}_i(x+\delta x)$ is "parallel" to $X_i(x)$. The difference $X_i(x+\delta x) - \mathfrak{X}_i(x+\delta x)$ is tensorially meaningful, and gives rise to the covariant derivative in the limit.*

We need to identify at $x+\delta x$ a "stand-in" for $X_i(x)$—a vector $\mathfrak{X}_i(x+\delta x)$ obtained by (in Weyl's phrase) the "parallel transport" of $X_i(x)$ from x to $x+\delta x$. With the aid of such an object we would construct

$$\frac{X_i(x+\delta x) - \mathfrak{X}_i(x+\delta x)}{\delta x}$$

which provides escape from our former transformational problem. Adopt this *definition* of the infinitesimial parallel transport process:

$$\mathcal{X}_i(x+\delta x) \equiv X_i(x) + X_a(x)\Gamma^a{}_{ij}(x)\delta x^j \tag{27}$$

We then have

$$\begin{aligned} X_i(x+\delta x) - \mathcal{X}_i(x+\delta x) &= \big\{X_i(x+\delta x) - X_i(x)\big\} - X_a(x)\Gamma^a{}_{ij}(x)\delta x^j \\ &= \big\{X_{i,j}(x) - X_a(x)\Gamma^a{}_{ij}(x)\big\}\delta x^j \\ &\equiv X_{i;j}(x)\delta x^j \end{aligned}$$

which gives back (24.1) and permits reconstruction of all that has gone before.

The parallel transport concept permits formulation of a valuable *metric-independent theory of geodesics*, which I now sketch. Let $\mathcal{C}$ refer as before to an arbitrarily parameterized[32] curve $x^i(t)$ which has been inscribed on $\mathcal{M}$. Indexed objects $v^i(t) \equiv \frac{d}{dt}x^i(t)$ are associated with the points of $\mathcal{C}$, and are readily seen to transform as contravariant vectors (essentially because the differentials dx^i do). It is natural to

call $v^i(t)$ the vector "tangent" to $\mathcal{C}$ at $x^i(t)$

Introduce

$$\mathcal{V}^i(t+\delta t) \equiv v^i(t) - v^a(t)\Gamma^i{}_{ab}(x(t))v^b(t)\delta t \tag{28}$$

to describe the result of parallel transporting $v^i(t)$ from $x^i(t)$ to the neighboring point $x^i(t) + v^i(t)\delta t$ and ask: How does $\mathcal{V}^i(t+\delta t)$ compare to $v^i(t+\delta t)$? If $\mathcal{C}$ is geodesic we expect those two to be parallel; i.e., we *expect geodesics to be generated by parallel transportation of a tangent*. One is tempted to look to the implications of $\mathcal{V}^i(t+\delta t) = v^i(t+\delta t)$ —i.e., of

$$\tfrac{d}{dt}v^i + \Gamma^i{}_{ab}\, v^a v^b = 0 \tag{29}$$

—but that equation is *not stable with respect to parametric regraduation*: an adjustment $t \to \tau = \tau(t)$ would cause (29) to become (see again (20), and agree in the present instance to write $w^i \equiv dx^i/d\tau = (d\tau/dt)^{-1}\, v^i = v^i/\dot\tau$)

$$\tfrac{d}{d\tau}w^i + \Gamma^i{}_{ab}\, w^a w^b = w^i\, \tfrac{d}{d\tau}\log\tfrac{dt}{d\tau} = -w^i\, \ddot\tau/\dot\tau^2$$

which is structurally distinct from (29). Evidently the most we can require is that $\mathcal{V}^i(t+\delta t) \sim v^i(t+\delta t)$; i.e., that there exists a $\varphi(t)$ such that

$$v^i(t) - v^a(t)\Gamma^i{}_{ab}(x(t))v^b(t)\delta t = [1 - \varphi(t)\delta t]\cdot v^i(t+\delta t)$$

Then in place of (29) we have

$$\tfrac{d}{dt}v^i + \Gamma^i{}_{ab}\, v^a v^b = v^i\cdot\varphi \tag{30.1}$$

which upon reparameterization $t \to \sigma = \sigma(t)$ becomes (write $u^i \equiv \frac{d}{d\sigma}x^i = v^i/\dot\sigma$)

$$\tfrac{d}{d\sigma}u^i + \Gamma^i{}_{ab}\, u^a u^b = u^i\cdot(\dot\sigma\phi - \ddot\sigma)/\dot\sigma^2 \tag{30.2}$$

[32] "Arc-length parameterization" is, in the absence of a metric, not an option.

where $\phi(\sigma(t)) \equiv \varphi(t)$. Notice that if we are given any instance of (30.1) then we have only to take $\sigma(t)$ to be any solution of $\dot{\sigma}\varphi - \ddot{\sigma} = 0$[33] to bring (30.2) to the form

$$\tfrac{d}{d\sigma}u^i + \Gamma^i{}_{ab}\, u^a u^b = 0 \tag{30.3}$$

so even in the absence of $g_{ij}(x)$ a "natural parameterization" is available, and when it is employed the geodesic condition (30.1) reads (30.3), which *does* possess the design which at (29) was originally conjectured.[34]

On *metrically* connected manifolds the specialization $\Gamma^i{}_{ab} \mapsto \left\{ {i \atop ab} \right\}$ becomes natural. It becomes natural, moreover, to adopt s-parameterization, and from $(ds)^2 = g_{ab}\, dx^a dx^b$ we conclude that

$$\text{the tangent vector } u^i(s) \equiv \tfrac{d}{ds}x^i(s) \text{ is a } \textit{unit} \text{ vector: } g_{ab}u^a u^b = 1$$

Equation (30.3) gives back (20.2), but with this added information: *The unit tangents to a Riemannian geodesic are parallel transports of one another*. The question which motivated this discussion—What has covariant differentiation to do with geodesic design?—has now an answer. But the concept of parallel tranport (introduced by Levi-Civita in 1917—too late to do Einstein any immediate good, but immediately put to general relativistic work by Weyl) has yet other things to teach us.

---

[33] The solution resulting from initial conditions $\sigma(0) = 0$, $\dot{\sigma}(0) = 1$ can be described

$$\sigma(t) = \int_0^t \exp\left\{ \int_0^{t'} \varphi(t'')\, dt'' \right\} dt' = t + \tfrac{1}{2}\varphi(0)t^2 + \cdots$$

[34] Given a tensor field $X^i(x)$, and a curve $\mathcal{C}$ inscribed by $x^i(t)$ on an affinely connected manifold $\mathcal{M}$, the "absolute derivative"

$$\tfrac{\delta}{\delta t}X^i \equiv \tfrac{d}{dt}X^i + \Gamma^i{}_{ab}\, X^a v^b$$

(here $v^i \equiv dx^i/dt$ and $\frac{d}{dt}X^i = X^i{}_{,a}v^a$) describes the rate at which X^i is seen to change as one progresses $t \to t + \delta t$ along the curve. It is easy to show that $\delta X^i/\delta t$ responds

- tensorially to recoordinatization $x \to y$, and
- by the chain rule $\frac{\delta}{\delta\tau} = \frac{dt}{d\tau}\frac{\delta}{\delta t}$ to reparameterization $t \to \tau$.

so statements of the form $\frac{\delta}{\delta t}X^i = 0$ are transformationally stable. Why, therefore, was (29), *not* stable? Because $X^i \mapsto v^i$ introduces an object which carries its own parameter-dependence. See "Non-Riemannian Spaces," the final chapter in J. L. Synge & A. Schild's *Tensor Calculus* (1952), for a detailed account of the absolute derivative and its applications. It was, by the way, from this text that I myself learned tensor calculus while an undergraduate, and in my view it remains the text best calculated to serve the needs of young physicists.

Curvature. Inscribe a closed curve $\mathcal{C}$ on the Euclidean plane. Parallel transport of a vector $\boldsymbol{X}$ around $\mathcal{C}$ (taken in Figure 5 to be a triangle) returns $\boldsymbol{X}$ to its

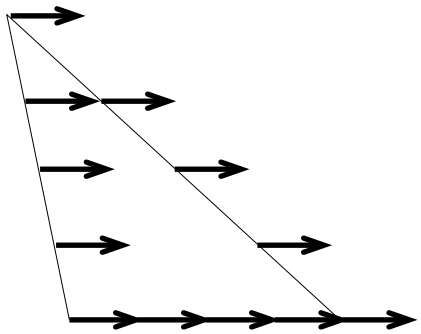


FIGURE 5: *Parallel transport of a vector along a closed path (here taken to be triangular) inscribed on the Euclidean plane. The vector returns home unchanged by its adventure.*

initial value. That this is, in general, *not* the case if $\mathcal{C}$ has been inscribed on a *curved* surface $\Sigma^2 \in E^3$ is illustrated in Figure 6. In that figure the apex of the spherical triangle sits at the pole ($\theta = 0$) and the equatorial base points differ in longitude by $\varphi = \phi_2 - \phi_1$; $\boldsymbol{X}$, upon return to its point of departure (where $\theta = \frac{\pi}{2}, \phi = \phi_1$), has been rotated $\circlearrowleft$ through angle φ. Spherical trigonometry supplies the information that

$$\text{(sum of interior angles)} - \pi \equiv \text{“spherical excess”} = \text{area}/R^2 \tag{31}$$

which acquires interest from the observation that

$$\text{“spherical excess”} = \varphi = \text{angle through which } \boldsymbol{X} \text{ is rotated}$$

This result—though special to the case illustrated—has a “generalizable look about it” ... and indeed: the “Gauss-Bonnet theorem” (which I will discuss in a moment: see Figure 6) asserts that

$$2\pi - \text{(sum of exterior angles)} - \int_{\mathcal{C}} \kappa_g \, ds = \iint_{\mathcal{D}} K \, dS \tag{32}$$

which in the case illustrated—where

- $\kappa_g = 0$ because $\mathcal{D}$ is bounded by geodesics;
- $K = 1/R^2$ everywhere because the surface is spherical

—gives back (31).

The meanings of the terms which enter into statement of the Gauss-Bonnet theorem are evident with one exception; I refer to the “geodesic curvature,” the meaning of which is explained in the caption to Figure 8.

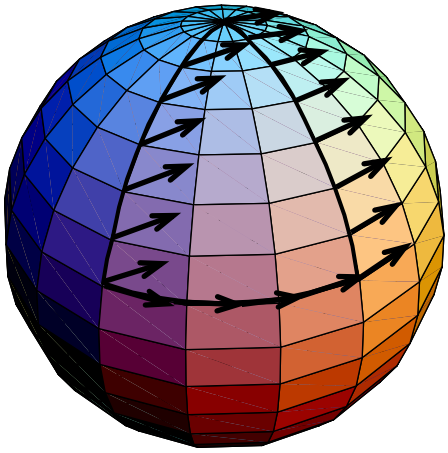


FIGURE 6: *Parallel transport of a vector along a triangular path inscribed on a sphere of radius R. The base points sit on the equator, the apex is at the pole, the sides are geodesic, and the circulation sense $\circlearrowleft$ has been counterclockwise.*

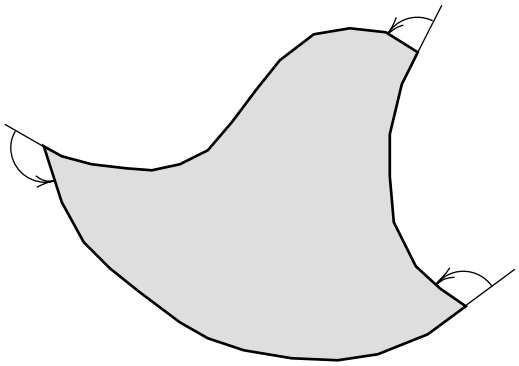


FIGURE 7: *The shaded region $\mathcal{D}$ is bounded by a closed contour $\mathcal{C}$ which has been inscribed on a Gaussian surface $\Sigma^2 \in E^3$. The Gauss-Bonnet theorem (32) refers to situations in which the number of vertices is arbitrary, the bounding arcs need not be geodesic, and the Gaussian curvature K may vary from point to point within $\mathcal{D}$.*

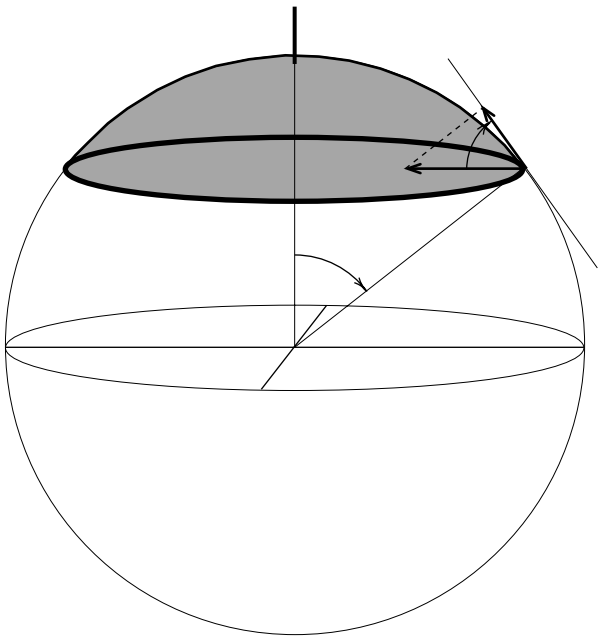


FIGURE 8: *The shaded cap $\mathcal{D}$, bounded by a circle of constant θ, has area $2\pi R^2(1-\cos\theta)$, and its boundary $\mathcal{C}$ presents no vertices. The Gauss-Bonnet theorem (32) therefore asserts that*

$$2\pi - \int_{\mathcal{C}} \kappa_g \, ds = 2\pi(1-\cos\theta) \qquad (i)$$

The curvature κ of the boundary $\mathcal{C}$—thought of as a space curve—arises from

$$\tfrac{d}{ds}\boldsymbol{x}(s) = \boldsymbol{T}(s) \quad \textit{and} \quad \tfrac{d}{ds}\boldsymbol{T}(s) = \kappa(s)\,\boldsymbol{n}(s)$$

(here $\boldsymbol{T}$ is the unit tangent to $\mathcal{C}$ at $\boldsymbol{x}(s)$, and $\boldsymbol{n}$ is the unit normal in the plane of $\mathcal{C}$), and in the present instance has constant value

$$\kappa = 1/(\textit{radius of curvature}) = 1/(R\sin\theta)$$

The "geodesic curvature" κ_g refers to the length of the projection of $\kappa\boldsymbol{n}$ onto the tangent plane, and is in the present instance given by

$$\kappa_g = \cos\theta/(R\sin\theta)$$

Therefore

$$\int_{\mathcal{C}} \kappa_g \, ds = \kappa_g \cdot 2\pi R\sin\theta = 2\pi\cos\theta$$

—in precise agreement with (i).

Generally, $\mathcal{C}$ is locally geodesic if $\boldsymbol{n}$ stands normal to the local tangent plane; it follows that

$$\kappa_g = 0 \quad \text{everywhere on a geodesic}$$

(whence the name), and that the geodesics inscribed on a Gaussian surface Σ^2 are "as uncurved as possible." The Gauss-Bonnet theorem (32), insofar as it entails

- exploration of the boundary $\mathcal{C} = \partial\mathcal{D}$ on the left
- exploration of the interior of $\mathcal{D}$ on the right,

bears a family resemblance to Stokes' theorem.[35] One expects it to be the case—as, indeed, it is—that κ_g can be described by operations intrinsic to the Gaussian surface.[36]

The circumstance discussed above—and illustrated in Figure 6 as it pertains to one particular Gaussian surface—was recognized by Riemann to pertain in metrically connected spaces of any dimension, and in fact it presumes only the affine connection needed to lend meaning to "parallel transport." The

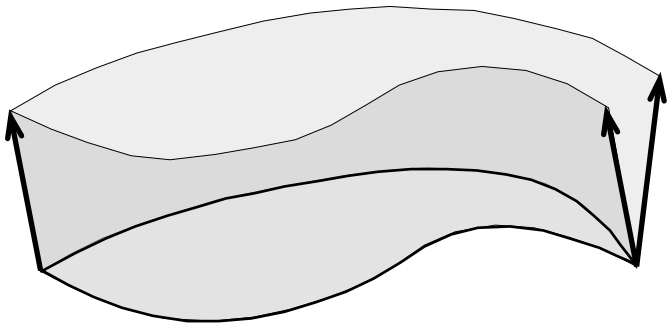


FIGURE 9: *The result of parallel transport from P to Q in on affinely connected manifold is typically path-dependent, and that fact can be taken to be the defining symptom of "curvature."*

basic phenomenon is illustrated in the preceding figure, and can be approached analytically in several ways. One might, for example, develop a formal description of the X^i_Q which results from parallel transport of X^i_P along $\mathcal{C}$, and then examine the δX^i_Q which results from variation of the path.[37] It is far easier and more efficient, however, to look to the δX^i which results from comparison of a pair of *differential paths* (Figure 10); one's interest is then

[35] The point is developed on pp. 45–48 of "Ellipsometry" (1999). The motivation there comes not from general relativity but from optics.

[36] For indication of how this can be accomplished, see the *Encyclopedic Dictionary of Mathematics* (1993), p. 1731.

[37] This is the program pursued on pp. 134–137 in notes previously cited.[28]

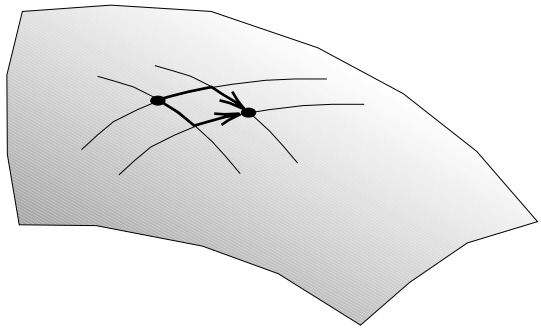


FIGURE 10: *Alternative paths* $\{\delta u \text{ then } \delta v\}$ *and* $\{\delta v \text{ then } \delta u\}$ *linking point* P *to a neighboring point in* $\mathcal{M}$.

directed to the expression

$$\begin{aligned}\delta X^i &= \Big\{\frac{\delta}{\delta v}\frac{\delta}{\delta u} - \frac{\delta}{\delta u}\frac{\delta}{\delta v}\Big\}X^i \delta u \delta v \\ &= \big\{X^i{}_{;jk} - X^i{}_{;kj}\big\}\frac{\partial x^j}{\partial u}\frac{\partial x^k}{\partial v}\delta u \delta v\end{aligned}$$

and thus by simple calculations to the statements

$$\begin{aligned}\big\{X^i{}_{;jk} - X^i{}_{;kj}\big\} &= (X^i{}_{,j} + X^a \Gamma^i{}_{aj})_{,k} + (X^b{}_{,j} + X^a \Gamma^b{}_{aj})\Gamma^i{}_{bk} \\ &\qquad - \text{ditto with } j \text{ and } k \text{ interchanged} \\ &= X^a{}_{,k}\Gamma^i{}_{aj} + X^b{}_{,j}\Gamma^i{}_{bk} + X^a(\Gamma^i{}_{aj,k} + \Gamma^b{}_{aj}\Gamma^i{}_{bk}) \\ &\qquad - \text{ditto with } j \text{ and } k \text{ interchanged} \\ &= -X^a R^i{}_{ajk} \qquad (33.1)\end{aligned}$$

$$R^i{}_{ajk} \equiv \Gamma^i{}_{ak,j} - \Gamma^i{}_{aj,k} + \Gamma^b{}_{ak}\Gamma^i{}_{bj} - \Gamma^b{}_{aj}\Gamma^i{}_{bk} \qquad (34)$$

$$\big\{X_{i;jk} - X_{i;kj}\big\} = +X_a R^a{}_{ijk} \qquad (33.2)$$

In the matrix notation $\mathbb{\Gamma}_i \equiv \|\Gamma^a{}_{bi}\|$ the definition (33) becomes

$$\mathbb{R}_{jk} = \partial_j \mathbb{\Gamma}_k - \partial_k \mathbb{\Gamma}_j + \mathbb{\Gamma}_j \mathbb{\Gamma}_k - \mathbb{\Gamma}_k \mathbb{\Gamma}_j \qquad (35)$$

It is an implication of the design of (33) that $R^i{}_{ajk}$ transforms tensorially (as a mixed fourth-rank tensor of zero weight), even though it has been assembled from objects which do *not* transform tensorially. On *metrically* connected manifolds it becomes natural in place of (34) to write

$$R^i{}_{ajk} = \frac{\partial}{\partial x^j}\left\{\begin{matrix} i \\ ak \end{matrix}\right\} - \frac{\partial}{\partial x^k}\left\{\begin{matrix} i \\ aj \end{matrix}\right\} + \left\{\begin{matrix} i \\ bj \end{matrix}\right\}\left\{\begin{matrix} b \\ ak \end{matrix}\right\} - \left\{\begin{matrix} i \\ bk \end{matrix}\right\}\left\{\begin{matrix} b \\ aj \end{matrix}\right\} \qquad (36)$$

This is the *Riemann curvature tensor*, which Riemann allegedly obtained

without benefit either of a theory of connections or of a tensor calculus.[38] The right side of (34) describes its (metric-independent) affine generalization.

From (34) it follows readily that

$$R^i{}_{ajk} = -R^i{}_{akj} \quad : \quad \text{antisymmetry in last subscripts} \tag{37.1}$$
$$R^i{}_{ajk} + R^i{}_{jka} + R^i{}_{kaj} = 0 \quad : \quad \text{cyclic symmetry} \tag{37.2}$$

On metrically connected manifolds we can use g_{ij} to construct

$$R_{iajk} \equiv g_{ib} R^b{}_{ajk} \tag{38}$$

which is found to possess, in addition to those symmetry properties, also two others:

$$R_{iajk} = -R_{aijk} \quad : \quad \text{antisymmetry in first subscripts} \tag{37.3}$$
$$R_{iajk} = +R_{jkia} \quad : \quad \text{symmetry in first/last pair-interchange} \tag{37.4}$$

Careful counting,[39] based upon those overlapping statements, shows that in the N-dimensional case R_{iajk} possesses a total of $\# \equiv \frac{1}{12}N^2(N^2-1)$ independent components; as N ranges on $\{2,3,4,5,\dots\}$ $\#$ ranges on $\{1,6,20,50,\dots\}$.

Direct computation establishes also that the Riemann tensor satisfies also a population of first derivative conditions which can be written

$$\mathbb{R}_{ij;k} + \mathbb{R}_{jk;i} + \mathbb{R}_{ki;j} = \mathbb{O} \tag{39}$$

and are called *Bianchi identities*.

It was remarked in connection with (22) that the replacement

$$\partial_j X_i \to \mathcal{D}_j X_i \equiv \partial_j X_i - X_a \Gamma^a{}_{ij}$$

"has much in common with the spirit of gauge field theory." I had then in mind the minimal coupling adjustment $\partial_\mu \psi \to \mathcal{D}_\mu \psi \equiv \partial_\mu \psi - ig\psi A_\mu$ which is basic to the latter theory. The parallel became more striking when it developed (compare (23) with (3–8)) that both the connection $\Gamma^a{}_{ij}$ and the gauge field A_μ had to *transform by acquisition of an additive term* if they were to perform their respective "compensator" roles. I draw attention now to the striking fact that (35) had a precise precursor in the equation

$$\boldsymbol{F}_{\mu\nu} \equiv (\partial_\mu \boldsymbol{A}_\nu - \partial_\nu \boldsymbol{A}_\mu) - ig(\boldsymbol{A}_\mu \boldsymbol{A}_\nu - \boldsymbol{A}_\nu \boldsymbol{A}_\mu)$$

[38] I say "allegedly" because I can find no hint of any such thing in Riemann's collected works! Eddington—I suspect with good historical cause—refers in his *Mathematical Theory of Relativity* (1923) always to the "Riemann-Christoffel tensor."

[39] See p. 86 in Synge & Schild.[34]

which at (3–101) served to define the non-Abelian analog of the electromagnetic field tensor. And the Bianchi identities (39) were anticipated at (1–126). Authors—writing in the presumption that their readers possess some familiarity with general relativity—most typically point to those formal parallels in an effort to make gauge field theory seem less alien.[40] My own intent has been the reverse. But pretty clearly, there must exist some sufficiently elevated viewpoint from which, in sufficiently fuzzy focus, gauge field theory and Riemannian geometry (if not general relativity itself) appear to be "the same thing."

Continuing in the presumption that $\mathcal{M}$ is metrically connected (i.e., that we are doing Riemannian geometry, and excluding from consideration the more relaxed structures contemplated by Weyl and others[34]), we observe that

$$g^{ab}R_{abij} = g^{ab}R_{ijab} = 0$$

and

$$g^{ab}R_{aijb} = -g^{ab}R_{aibj} = -g^{ab}R_{iajb} = g^{ab}R_{iabj}$$

are consequences of (37); there is, in other words, only one way (up to a sign) to contract the metric into the curvature tensor. The *Ricci tensor* is defined[41]

$$R_{ij} \equiv g^{ab}R_{aijb} = R^{a}{}_{ija} \tag{40}$$

and is, by (37), symmetric. More obviously, there is only one way to contract the metric into the Ricci tensor; it yields the *curvature invariant*:

$$R \equiv g^{ij}R_{ij} = R^{i}{}_{i} \tag{41}$$

The Bianchi identities (39), with input from (37), can be written

$$R_{abij;k} - R_{bajk;i} - R_{abik;j} = 0 \tag{42.1}$$

and, when contracted into $g^{ai}g^{bj}$, give $R_{;k} - 2R^{i}{}_{k;i} = 0$ which can be expressed

$$(R^{i}{}_{k} - \tfrac{1}{2}R\delta^{i}{}_{k})_{;i} = 0 \tag{42.2}$$

The preceding equations, known as the "contracted Bianchi identities," can be read as an assertion that the (covariant) divergence of the *Einstein tensor*

$$G_{ij} \equiv R_{ij} - \tfrac{1}{2}Rg_{ij} \quad : \quad G_{ij} = G_{ji} \tag{43}$$

---

[40] See, for example, p. 638 in M. Kaku, *Quantum Field Theory* (1993), or §15.1 "The geometry of gauge invariance" in M. E. Peskin & D. V. Schroeder, *An Introduction to Quantum Field Theory* (1995).

[41] Careful! I have followed the conventions of Synge & Schild, but many modern authors write $R^{i}{}_{ajk}$ where I have written $R^{i}{}_{akj}$; i.e., they work with the *negative* of my Riemann tensor. But where I write $R_{ij} \equiv R^{a}{}_{ija}$ they write $R_{ij} \equiv R^{a}{}_{iaj}$, so we are in agreement on the definition of the Ricci tensor, and of R.

vanishes: $G^{ij}{}_{;i} = 0$. The contracted identities (42.2) were first noted by Aurel Voss in 1880, rediscovered by Ricci in 1889 and rediscovered again by Luigi Bianchi in 1902, the argument in each instance being heavily computational; the observation that (42.2) follows quickly from "the" Bianchi identities (42.1) was not made until 1922. The geometrical statements (42.2) strike the eye of a physicist as a quartet of *conservation laws*, and it is as such that they have come to play an important role in gravitational theory. But they remained unknown to Einstein and Hilbert (and to most other experts, except for Weyl) during the years when they were most needed.[42]

But what have the Riemann curvature tensor and its relatives got to do with "curvature" in any familiar geometrical sense? To start with the almost obvious: *If* there exists *some* coordinate system x with respect to which the components of g_{ij} *all become constant*, then (see again the definitions (26) of the Christoffel symbols) the $\left\{ {i \atop jk} \right\}$ all *vanish* in those coordinates, and so also therefore (see again the definition (36)) do the components $R^a{}_{ijk}$ of the curvature tensor. Coordinate adjustment $x \to y$ will, in general, cause the metric to no longer be constant, and the Christoffel symbols to no longer vanish, but because $R^a{}_{ijk}$ transforms *as a tensor* it will *vanish in all coordinate systems*. It turns out that the converse is also true:

$$R^a{}_{ijk} = 0 \text{ if and only if there exists a coordinate system in which the metric } g_{ij} \text{ becomes constant}$$

In such a coordinate system g_{ij} can, by rotation, be diagonalized, and by dilation one can arrange to have only ± 1's appear on the principal diagonal. If all signs are positive, then one has, in effect, "Cartesian coordinatized Euclidean N-space," and in all cases it becomes sensible to interpret $R^a{}_{ijk} = 0$ as a "flatness condition." Which is to interpret $R^a{}_{ijk} \neq 0$ as the condition that the Riemannian manifold $\mathcal{M}$ be "not flat, or curved." To test the plausibility of the interpretation we must look to circumstances in which we have some *prior* conception of curvature. Though we might look to "spheres in N-space,"[43] concerning which we have some sense of what the constant curvature should be, it makes more sense (and is easier) to look to the case $N = 2$, where our intuitions are vivid, and where additionally we have Gauss' analytical accomplishments to guide us.

In the case $N = 2$ the Riemann tensor has only one independent element;

$$R_{abij} = \begin{cases} \pm R_{1212} & \\ \text{else } 0 & \end{cases} \quad \text{on 2-dimensional manifolds}$$

Look to the case $(ds)^2 = r^2(d\theta)^2 + (r\sin\theta)^2(d\phi)^2$, which we saw at (14) refers to the spherical coordinatization of a globe of radius r (Figure 3). In *MTW* (at p. 340 in Chapter 14: "Calculation of Curvature") we are walked through

[42] See §15c in Pais for further historical commentary.

[43] See §4 in "Algebraic theory of spherical harmonics" (1996) for indication of how this might be done.

the demonstration that

- $\Gamma^{\theta}{}_{\phi\phi} = -\sin\theta\cos\theta$; $\Gamma^{\phi}{}_{\phi\theta} = \Gamma^{\phi}{}_{\theta\phi} = \cot\theta$; other Γ's vanish;
- $R_{\theta\phi\theta\phi} = r^2\sin^2\theta$, with other elements determined by symmetries;
- $R^{\theta}{}_{\theta} = R^{\phi}{}_{\phi} = 1/r^2$; $R^{\theta}{}_{\phi} = R^{\phi}{}_{\theta} = 0$;
- $R = 2/r^2$

So at least in this hallowed case the curvature tensor speaks (through the curvature scalar R) directly to what we are prepared to call the "curvature of the sphere." More generally,[44] the Gaussian curvature can be described

$$K = \frac{R_{1212}}{g}$$
$$= \frac{r^2\sin^2}{r^4\sin^2\theta} = 1/r^2 \quad \text{in the spherical case just considered}$$

We conclude on this evidence that $R^{a}{}_{ijk}$ reproduces what we already knew about curvature, and puts us in position to say things we didn't already know. More particularly, it achieves a vast generalization of ideas pioneered by Gauss —ideas to which the young Einstein had been exposed, but had paid no special attention.

Variational approach to the gravitational field equations. Once Einstein had acquired the conviction that

- gravitation is an artifact of the (Riemannian) geometry of spacetime, and
- general covariance is an essential formal property of any theory of gravity

he set out discover the field equations to which $g_{\mu\nu}$ must be subject.[45] His principal guidance was supplied by Newton; i.e., by Poisson's equation

$$\nabla^2\varphi = 4\pi G\rho \tag{5}$$

from which he inferred that $g_{\mu\nu}(x)$ should satisfy a generally covariant coupled system of second-order partial differential equations—equations which are, in particular, linear and homogeneous in the second partials.[46] The tensor analytic theory of Riemannian manifolds supplies, as we have seen, only limited material, and led him to contemplate field equations of the narrowly constrained design

$$R_{\mu\nu} + \alpha R g_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu} \tag{44}$$

where α and Λ are adjustable constants, $T_{\mu\nu}$ is the (necessarily symmetric) stress-energy tensor of such non-gravitational (material, electromagnetic) fields as may inhabit spacetime, and κ is to be fixed by dimensional considerations (units). It was to achieve the conservation law (42.2) that he—somewhat

[44] See McCleary,[20] p. 155 for the demonstration.

[45] I revert to Greek indices to emphasize that we now inhabit not some arbitrary Riemannian manifold, but the *physical spacetime manifold*, which is understood to be 4-dimensional.

[46] See his *The Meaning of Relativity* (3rd revised edition 1950), p. 84.

tentatively—set $\alpha = -\frac{1}{2}$ and Λ (the so-called "cosmological constant") equal to zero.

Hilbert, on the other hand, chose to look upon the problem posed by Einstein as a straightforward problem in Lagrangian field theory; i.e., to work from the assumption that the field equations can be derived from a variational principle.[47] His initial effort, therefore, was to build Einstein's requirements into the design of a Lagrange density. Before we attempt to follow in Hilbert's steps, let us...

Recall again (from elementary calculus) that integrals respond to a change of variables $x \to y$ by "picking up a Jacobian"

$$\iiiint_{\text{bubble}} f(x)\, dx^0 dx^1 dx^2 dx^3$$
$$= \iiiint_{\text{image bubble}} \underbrace{f(x(y)) \left| \frac{\partial(x^0, x^1, x^2, x^3)}{\partial(y^0, y^1, y^2, y^3)} \right|}_{F(y)} dy^0 dy^1 dy^2 dy^3$$

The general covariance of an action functional

$$S \equiv \iiiint_{\mathcal{R}} \mathfrak{L}\ dx^0 dx^1 dx^2 dx^3 = \iiiint_{\mathcal{R}'} \mathfrak{L}'\ dx'^0 dx'^1 dx'^2 dx'^3$$

therefore requires that $\mathfrak{L}$ respond to $x \to x'$ by the rule

$$\mathfrak{L} \to \mathfrak{L}' = \mathfrak{L} \cdot |\tfrac{\partial x}{\partial x'}|$$

i.e., as a *scalar density of unit weight*. In special relativity the Jacobian has value ± 1, so (as was remarked already at p. 9 in Chapter 2) to draw attention to the point just made is to underscore a "distinction without a difference." But in general relativity the point assumes non-trivial importance. It was remarked previously that in metric theories $\sqrt{g}$ supplies the "prototypical instance" of a scalar density of unit weight. We expect therefore to be able to write

$$\mathfrak{L} = \sqrt{g} \cdot (\text{generally covariant scalar of zero weight})$$

Another preparatory detail. Assume for the moment that $[x^\mu] = \textit{length}$ for all μ. Then the $g_{\mu\nu}$ are all dimensionless, giving[48]

$$[R_{aijk}] = [R_{ij}] = [R] = 1/(\text{length})^2$$

[47] Pauli, writing in *1921*, considered reliance upon a variational principle to be a defect, "unacceptable to physicists"...as at the time it assuredly was; today most physicists would consider this to be the principal virtue of Hilbert's approach.

[48] The following statements remain in force even if some of the generalized coordinates are (for example) angles, since compensating dimensionality—sufficient to preserve $[ds] = \textit{length}$ —will attach then to associated components of the metric. For that same reason, $[\sqrt{g}\, dx^0 dx^1 dx^2 dx^3] = (\textit{length})^4$ in all cases.

We want to achieve $[\mathfrak{L}] = \textit{energy density} = ML^{-1}T^{-2}$, and have only R, c (a velocity) and G (of dimension $M^{-1}L^3T^{-2}$) to work with. Write

$$\mathfrak{L} \sim G^\alpha c^\beta R^\sigma \quad : \quad \begin{aligned} ML^{-1}T^{-2} &= (M^{-1}L^3T^{-2})^\alpha (LT^{-1})^\beta L^{-2\sigma} \\ &= M^{-\alpha}L^{3\alpha+\beta-2\sigma}T^{-2\alpha-\beta} \end{aligned}$$

and find that necessarily $\alpha = -1$, $\beta = 4$ and $\sigma = 1$. We expect therefore to have

$$\mathfrak{L} = (\text{dimensionless numeric})(c^4/G)\cdot\sqrt{g}\,\big\{R - 2\Lambda\big\} \tag{45}$$

where Λ is a dimensioned constant ($[\Lambda] = [R]$) and the anticipates recovery of (44). If we possessed a "natural length" λ then we could install also a factor of the form

$$\big\{1 + a(\lambda^2 R) + b(\lambda^2 R)^2 + \cdots\big\}$$

but (G,c)-theory presents no such object; $(G,c,\hbar)$-theory, on the other hand, does: we have the

$$\text{Planck length } \lambda \equiv \hbar G/c^3 = 1.616 \times 10^{-33}\,\text{cm}$$

which is too small to be of any direct relevance to macroscopic theory.[49] The prefactor in (45) will come into play when we undertake to establish detailed contact with Poisson's equation (5); until then we will ignore it, writing[50]

$$\mathfrak{L} \sim \sqrt{g}\,\big\{R - 2\Lambda\big\} \tag{46}$$

In the same sense that (say) a sphere is "locally Euclidean" in the neighborhood of every point, we expect the spacetime manifold $\mathcal{M}$ to be "locally Lorentzian." More specifically, we expect to be able to write

$$\|g_{\mu\nu}(x)\| = \mathbb{A}^{\intercal}(x)\,\mathbb{G}\,\mathbb{A}(x) \quad \text{with} \quad \mathbb{G} \equiv \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{47.1}$$

with $\mathbb{A}(x)$ determined only to within a local Lorentz transformation:

$$\mathbb{A}(x) \to \mathbb{A}'(x) \equiv \mathbb{L}\,\mathbb{A}(x) \quad : \quad \mathbb{L}^{\intercal}\mathbb{G}\,\mathbb{L} = \mathbb{G} \tag{47.2}$$

---

[49] In this connection see S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (1972), p. 365.

[50] In four dimensions—exceptionally—it is possible to contemplate a scalar density of the design $\epsilon^{\alpha\mu\nu\sigma}R_{\alpha\mu\nu\sigma}$. But antisymmetry properties of the Levi-Civita tensor can be used to write

$$\begin{aligned} \epsilon^{\alpha\mu\nu\sigma}R_{\alpha\mu\nu\sigma} &= \tfrac{1}{3}\epsilon^{\alpha\mu\nu\sigma}(R_{\alpha\mu\nu\sigma} + R_{\alpha\nu\sigma\mu} + R_{\alpha\sigma\mu\nu}) \\ &= 0 \quad \text{by (37.2)} \end{aligned}$$

so such a scalar can play no role in the theory.

The implication is that we expect to have

$$g(x) \equiv \det \|g_{\mu\nu}(x)\| = -(\det \mathbb{A})^2 < 0 \tag{48}$$

It is in light of this circumstance, and to preserve manifest reality, that the general relativistic literature bristles with $\sqrt{-g}$ factors. I find the minus signs a distraction, so will (in the good company of Weinberg) drop them (i.e., I will write $\sqrt{-g} = i\sqrt{g}$ and absorb the i into the meaning of "$\sim$"), and pick them up again only when their presence makes a difference.

If we interpret the field functions to be $g_{\mu\nu}(x)$, and recall from (26) the definitions of the Christoffel symbols, then (46) assumes the design

$$\mathfrak{L}(g, \partial g, \partial\partial g)$$

which, owing to the presence of the second derivatives, requires—or appears to require—an extension of Lagrangian field theory.[51] Several familiar tricks are available: we might, for example, expand the number fields, writing $\mathfrak{L}(g, h, \partial h)$ ... though this, so far as I am aware, is never done. Alternatively, one can borrow a trick from Procca field theory (a trick which is often useful also in electrodynamics and in many other applications): we found at (2–31) that it is formally advantageous to consider U^μ and $G^{\mu\nu} \equiv \partial^\mu U^\nu - \partial^\nu U^\mu$ to be *independent* fields, even though it is clearly impossible to vary $G^{\mu\nu}$ while holding U^μ constant. "Palatini's method"[52] proceeds similarly: one opts to look upon $g_{\mu\nu} = g_{\nu\mu}$ and $\Gamma^\alpha{}_{\mu\nu} = \Gamma^\alpha{}_{\nu\mu}$ as formally independent fields (10 of the former, 40 of the latter), and hopes to recover the definitions (26.2) as forced implications of an expanded set of field equations

$$\left\{\partial_\sigma \frac{\partial}{\partial g_{\mu\nu,\sigma}} - \frac{\partial}{\partial g_{\mu\nu}}\right\}\mathfrak{L} = 0 \tag{49.1}$$

$$\left\{\partial_\sigma \frac{\partial}{\partial \Gamma^\rho{}_{\mu\nu,\sigma}} - \frac{\partial}{\partial \Gamma^\rho{}_{\mu\nu}}\right\}\mathfrak{L} = 0 \tag{49.2}$$

where the Lagrangian has now the design

$$\mathfrak{L}(g_{..}, \Gamma^{\cdot}{}_{..}, \partial\Gamma^{\cdot}{}_{..}) \sim \sqrt{g}\left\{g^{\alpha\beta}R_{\alpha\beta}(\Gamma^{\cdot}{}_{..}, \partial\Gamma^{\cdot}{}_{..}) - 2\Lambda\right\} \tag{50}$$

We look first to (49.1), where the absence from $\mathfrak{L}$ of any ∂g-dependence results in some welcome simplification. We have

$$\sqrt{g}\,R_{\alpha\beta}\frac{\partial}{\partial g_{\mu\nu}}g^{\alpha\beta} + \{R - 2\Lambda\}\frac{\partial}{\partial g_{\mu\nu}}\sqrt{g} = 0$$

and—drawing upon the soon-to-be-established information that

$$\frac{\partial}{\partial g_{\mu\nu}}g^{\alpha\beta} = -g^{\mu\alpha}g^{\beta\nu} \quad \text{and} \quad \frac{\partial}{\partial g_{\mu\nu}}\sqrt{g} = \tfrac{1}{2}\sqrt{g}\,g^{\mu\nu} \tag{51}$$

---

[51] A clever way to circumvent this problem—due to Dirac—will be described later.

[52] A. Palatini, "Deduzione invariantiva delle equazioni gravitazioni dal principio di Hamilton," Rend. Circ. Mat. Palermo **43**, 203 (1919); A. Einstein, "Einheitliche Feldtheorie von Gravitation und Elektrizität," Preussische Akademie der Wissenschaften (1925), p. 414.

—obtain $\sqrt{g}\left\{ -R^{\mu\nu} + \frac{1}{2}(R - 2\Lambda)g^{\mu\nu}\right\} = 0$, or again

$$R^{\mu\nu} - \tfrac{1}{2}Rg^{\mu\nu} + \Lambda g^{\mu\nu} = 0 \tag{52}$$

This is an "empty universe" instance ($T^{\mu\nu} = 0$) of Einstein's gravitational field equations (44). Einstein's tentative $\alpha = -\frac{1}{2}$ is seen to have been *forced by the variational principle*. Notice that the $\frac{1}{2}$ has its origin in the $\sqrt{\ }$; i.e., in the circumstance that $\mathfrak{L}$ transforms as a scalar *density*. It is interesting in the light of more recent developments that Einstein's $\Lambda = 0$ is *not* forced.[53]

I digress to establish equations (51). Observe first that differentiation of $x^{-1}x = 1$ gives $x\frac{d}{dx}x^{-1} + x^{-1} = 0$ and provides a demonstration (as if one were needed) that $\frac{d}{dx}x^{-1} = -x^{-2}$. Similarly... let $\mathbb{A} \equiv \|a_{ij}\|$ be an invertible square matrix, and let $\mathbb{A}^{-1}$ be notated $\|a^{ij}\|$. Differentiation of $a^{ik}a_{kj} = \delta^i{}_j$ gives

$$\frac{\partial a^{ik}}{\partial a_{pq}}a_{kj} + a^{ik}\frac{\partial a_{kj}}{\partial a_{pq}} = \frac{\partial a^{ik}}{\partial a_{pq}}a_{kj} + a^{ik}\delta^p{}_k\delta^q{}_j = 0$$

$$\frac{\partial a^{ik}}{\partial a_{pq}}a_{kh} = -a^{ip}\delta^q{}_h$$

Multiplication by a^{hj} gives

$$\frac{\partial a^{ij}}{\partial a_{pq}} = -a^{ip}a^{qj}$$

which establishes the first part of (51). To gain insight into the second part, suppose that $\mathbb{A}$ were 2×2; we would then have

$$a \equiv \det \mathbb{A} = a_{11}a_{22} - a_{12}a_{21}$$

and

$$\mathbb{A}^{-1} = a^{-1}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11}\end{pmatrix} = \begin{pmatrix} \frac{1}{a}\frac{\partial a}{\partial a_{11}} & \frac{1}{a}\frac{\partial a}{\partial a_{21}} \\ \\ \frac{1}{a}\frac{\partial a}{\partial a_{12}} & \frac{1}{a}\frac{\partial a}{\partial a_{22}}\end{pmatrix}$$

giving

$$a^{ij} = \frac{1}{a}\frac{\partial a}{\partial a_{ji}} = \frac{\partial}{\partial a_{ji}}\log\det\mathbb{A} \tag{53}$$

[53] The "cosmological term" $\Lambda g^{\mu\nu}$ was introduced into (52) by Einstein in 1917, in an effort to make general relativity conform to what he imagined to be the steady state of the universe; the event took place at equation (13a) in a paper "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie" of which an English translation can be found in the Dover edition of *The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity*. He had become disenchanted with the term already by 1923, and officially abandoned it in 1931 (see Pais' §15.e). For indication of why there is renewed interest in the term, see Alan Guth, *The Inflationary Universe* (1997), p. 283.

whence

$$\frac{\partial}{\partial g_{\mu\nu}}\sqrt{g} = \frac{1}{2\sqrt{g}}\frac{\partial}{\partial g_{\mu\nu}}g = \frac{1}{2\sqrt{g}}\, g\, g^{\nu\mu}$$

from which (in the 2-dimensional case) the desired result follows by $g^{\nu\mu} = g^{\mu\nu}$. To establish the general validity of (53) we use the Laplace expansion

$$a = a_{j1}A_{j1} + a_{j2}A_{j2} + \cdots + a_{jN}A_{jN}$$

to obtain

$$\frac{\partial a}{\partial a_{ji}} = A_{ji} \equiv \text{cofactor of } a_{ji}$$

But by Cramer's Rule

$$a^{ij} = a^{-1}A_{ji}$$

which completes the argument. We will have need in a moment also of the second of these corollary of (51):

$$(\sqrt{g}\,)_{,\sigma} = \tfrac{1}{2}\sqrt{g}\, g^{\mu\nu}g_{\mu\nu,\sigma} = -\tfrac{1}{2}\sqrt{g}\, g_{\mu\nu}g^{\mu\nu}{}_{,\sigma} \tag{54}$$

Look now to (49.2), the second set of "field equations," which can be written

$$\left\{\partial_\sigma \frac{\partial}{\partial \Gamma^\rho{}_{\mu\nu,\sigma}} - \frac{\partial}{\partial \Gamma^\rho{}_{\mu\nu}}\right\}\sqrt{g}\, g^{\alpha\beta}R_{\alpha\beta} = 0$$

with[54]

$$R_{\alpha\beta} = \Gamma^i{}_{\alpha k,\beta}\delta^k{}_i - \Gamma^i{}_{\alpha\beta,k}\delta^k{}_i + \Gamma^i{}_{\alpha j}\Gamma^j{}_{i\beta} - \Gamma^i{}_{\alpha\beta}\Gamma^j{}_{ik}\delta^k{}_j$$

Carefully performing the indicated differentiations,[55] we obtain

$$\begin{aligned}
&\partial_\sigma\Big\{\sqrt{g}\,\big[\tfrac{1}{2}g^{\mu\sigma}\delta^\nu{}_\rho + \tfrac{1}{2}g^{\nu\sigma}\delta^\mu{}_\rho - g^{\mu\nu}\delta^\sigma{}_\rho\big]\Big\} \\
&\quad = \sqrt{g}\,\big[g^{\mu\alpha}\Gamma^\nu{}_{\rho\alpha} + g^{\nu\alpha}\Gamma^\mu{}_{\rho\alpha} - g^{\mu\nu}\Gamma^\alpha{}_{\rho\alpha} - \tfrac{1}{2}g^{\alpha\beta}\Gamma^\mu{}_{\alpha\beta}\delta^\nu{}_\rho - \tfrac{1}{2}g^{\alpha\beta}\Gamma^\nu{}_{\alpha\beta}\delta^\mu{}_\rho\big]
\end{aligned}$$

which by reorganization (I make free use of the $\mu\nu$-symmetry of $g_{\mu\nu}$ and $\Gamma^\alpha{}_{\mu\nu}$, and have underscored a couple of places where I have introduced a term promptly to subtract it again) becomes

$$\begin{aligned}
&\big[\sqrt{g}\, g^{\mu\nu}\big]_{,\rho} + \sqrt{g}\big[g^{\alpha\nu}\Gamma^\mu{}_{\alpha\rho} + g^{\mu\alpha}\Gamma^\nu{}_{\alpha\rho} - g^{\mu\nu}\Gamma^\alpha{}_{\alpha\rho}\big] \\
&\qquad = \quad \tfrac{1}{2}\Big\{\big[\sqrt{g}\, g^{\mu\sigma}\big]_{,\sigma} + \sqrt{g}\,\big[g^{\alpha\sigma}\Gamma^\mu{}_{\alpha\sigma} + \underline{g^{\mu\alpha}\Gamma^\sigma{}_{\alpha\sigma} - g^{\mu\sigma}\Gamma^\alpha{}_{\alpha\sigma}}\,\big]\Big\}\delta^\nu{}_\rho \\
&\qquad\quad + \tfrac{1}{2}\Big\{\big[\sqrt{g}\, g^{\nu\sigma}\big]_{,\sigma} + \sqrt{g}\,\big[g^{\alpha\sigma}\Gamma^\nu{}_{\alpha\sigma} + \underline{g^{\nu\alpha}\Gamma^\sigma{}_{\alpha\sigma} - g^{\nu\sigma}\Gamma^\alpha{}_{\alpha\sigma}}\,\big]\Big\}\delta^\mu{}_\rho
\end{aligned}$$

[54] Depleted ranks here force me to press some Roman soldiers into Greek service. The δ's have been introduced to prevent repeated indices from appearing on any Γ, which simplifies calculation in the present context.

[55] The assumed symmetry of $\Gamma^\alpha{}_{\mu\nu}$ presents a formal problem similar to that confronted/resolved in connection with (1–44). The equations which follow have been written in such a way as to display *manifest $\mu\nu$-symmetry*.

Recalling from (25) some defining properties of the covariant derivative, we find that the preceding equation can be expressed

$$\left[\sqrt{g}\,g^{\mu\nu}\right]_{;\rho} = \tfrac{1}{2}\left[\sqrt{g}\,g^{\mu\sigma}\right]_{;\sigma}\delta^{\nu}{}_{\rho} + \tfrac{1}{2}\left[\sqrt{g}\,g^{\nu\sigma}\right]_{;\sigma}\delta^{\mu}{}_{\rho}$$

or again (here I adopt the notation of *CTW*, p. 502)

$$\mathfrak{g}^{\mu\nu}{}_{;\rho} - \tfrac{1}{2}\delta^{\mu}{}_{\rho}\,\mathfrak{g}^{\nu\sigma}{}_{;\sigma} - \tfrac{1}{2}\delta^{\nu}{}_{\rho}\,\mathfrak{g}^{\mu\sigma}{}_{;\sigma} = 0 \tag{55}$$

where $\mathfrak{g}^{\mu\nu} \equiv \sqrt{g}\,g^{\mu\nu}$ defines a "metric density" or unit weight. Look upon (55) as a homogeneous linear system of 40 equations in 40 unknowns:

$$\begin{pmatrix} \bullet & \bullet & \cdots & \bullet \\ \bullet & \bullet & \cdots & \bullet \\ \vdots & \vdots & & \vdots \\ \bullet & \bullet & \cdots & \bullet \end{pmatrix} \begin{pmatrix} \mathfrak{g}^{00}{}_{;0} \\ \mathfrak{g}^{01}{}_{;0} \\ \vdots \\ \mathfrak{g}^{44}{}_{;4} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Writing the 40×40 matrix into *Mathematica* we discover it to be non-singular, so (55) will be satisfied if and only if

$$\mathfrak{g}^{\mu\nu}{}_{;\sigma} = 0 \tag{56}$$

It remains (*i*) to show that this result implies (and is implied by) $g_{\mu\nu;\sigma} = 0$, and (*ii*) to discuss the remarkable significance of that fact. When written out in detail, (56) reads

$$(\sqrt{g}g^{\mu\nu})_{,\sigma} + \sqrt{g}g^{\alpha\nu}\Gamma^{\mu}{}_{\alpha\sigma} + \sqrt{g}g^{\mu\alpha}\Gamma^{\nu}{}_{\alpha\sigma} - \sqrt{g}g^{\mu\nu}\Gamma^{\alpha}{}_{\alpha\sigma} = 0$$

which when contracted into $g_{\mu\nu}$ gives

$$(\sqrt{g}\,)_{,\sigma}N + \sqrt{g}g_{\mu\nu}g^{\mu\nu}{}_{,\sigma} + \sqrt{g}\Gamma^{\alpha}{}_{\alpha\sigma} + \sqrt{g}\Gamma^{\alpha}{}_{\alpha\sigma} - N\sqrt{g}\Gamma^{\alpha}{}_{\alpha\sigma} = 0$$

where $N = \delta^{\alpha}{}_{\alpha}$ = dimension of the spacetime manifold. Drawing now upon (54), we have

$$(N-2)\Big\{(\sqrt{g}\,)_{,\sigma} - \sqrt{g}\Gamma^{\alpha}{}_{\alpha\sigma}\Big\} = 0$$

which by (25.2) becomes $(N-2)(\sqrt{g}\,)_{;\sigma} = 0$ and (if $N \neq 2$) entails

$$(\sqrt{g}\,)_{;\sigma} = 0 \tag{57}$$

This result, we not in passing, supplies the following often-useful information:

$$\Gamma^{\alpha}{}_{\alpha\sigma} = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^{\sigma}}\sqrt{g} = \frac{\partial}{\partial x^{\sigma}}\log\sqrt{g} \tag{58}$$

Now construct $\mathfrak{g}^{\alpha\beta}g_{\beta\nu} = \sqrt{g}\,\delta^{\alpha}{}_{\nu}$ and covariantly differentiate:

$$\mathfrak{g}^{\alpha\beta}{}_{;\sigma}g_{\beta\nu} + \mathfrak{g}^{\alpha\beta}g_{\beta\nu;\sigma} = (\sqrt{g}\,)_{;\sigma}\delta^{\alpha}{}_{\nu} + \sqrt{g}\,\delta^{\alpha}{}_{\nu;\sigma}$$

$$\Downarrow$$

$$\mathfrak{g}^{\alpha\beta}g_{\beta\nu;\sigma} = 0$$

Contract into $g_{\mu\alpha}$ and obtain

$$g_{\mu\nu;\sigma} = 0 \tag{59}$$

Clearly, the order of the argument could be reversed: (56) $\Leftrightarrow$ (59).

At (59) we have recovered—now as "field equations"—the conditions which were previously seen to be necessary and sufficient for $\Gamma^{\cdot}{}_{\cdot\cdot}$-mediated covariant differentiation and $g_{\cdot\cdot}$-mediated index manipulation to performable in either order, conditions which at (26.2) were seen to entail

$$\Gamma^{\sigma}{}_{\mu\nu} = \tfrac{1}{2}g^{\sigma\alpha}\{g_{\alpha\mu,\nu} + g_{\alpha\nu,\mu} - g_{\mu\nu,\alpha}\}$$

That system of equations serves to locate "Riemannian geometry" within the broader class of (torsion-free) "affine geometries," and makes precise the sense in which (in Riemannian geometry) *metric structure dictates affine structure*. Einstein, having been led to embrace the Principle of General Covariance, found (actually, Hilbert found) the design (46) of the Lagrangian $\mathfrak{L}$ to be essentially forced, but while he

- needed $\Gamma^{\sigma}{}_{\mu\nu}$ to construct $R^{\sigma}{}_{\mu\nu\rho}$
- needed $g_{\mu\nu}$ to construct $R \equiv g^{\sigma\rho}g^{\mu\nu}R^{\sigma}{}_{\mu\nu\rho}$ and to supply a density

he did not need to *assume the compatability* of those connections; compatability was (at least within the Palatini formalism) *automatically enforced by the variational principle*.

To describe the geometry of spacetime in a world *not* devoid of matter Einstein (Hilbert) makes the *ad hoc* adjustment

$$\mathfrak{L}_{\text{gravitational field}} \quad \longmapsto \quad \mathfrak{L}_{\text{gravitational field}} + \mathfrak{L}_{\text{matter}} \tag{60}$$

Note the absence of an explicit "interaction term." The coupling of matter to gravitation is accomplished implicitly, through in the requirement that $\mathfrak{L}_{\text{matter}}$ be a *generally covariant density* (i.e., by introducing occasional $\sqrt{g}$-factors, and replacing some commas with semi-colons!). In the simplest case one might write (compare (2–11))

$$\mathfrak{L}_{\text{matter}} \sim \sqrt{g}\{g^{\mu\nu}\varphi_{;\mu}\varphi_{;\nu} - \varkappa^2\varphi^2\}$$

on the assumption that φ is a weightless scalar field (in which case $\varphi_{;\mu} = \varphi_{,\mu}$), or still more simply

$$\mathfrak{L}_{\text{matter}} \sim \{g^{\mu\nu}\varphi_{;\mu}\varphi_{;\nu} - \varkappa^2\varphi^2\}$$

on the assumption that φ transforms as a scalar density of weight $w = \frac{1}{2}$ (in which case $\varphi_{;\mu} = \varphi_{,\mu} - \frac{1}{2}\varphi\Gamma^{\alpha}{}_{\alpha\mu}$). I will return later to discussion of some of the general relativistic ramifications of (60).

I have organized the preceding discussion in a way intended to emphasize that gravitational field theory is (in at least its variational aspects) *classical field theory like any other*. All followed from (49). I cannot account for the fact that the authors of the standard monographs prefer *not* to work from those elegant equations, but to "reinvent" variational methodology as they go along.[56] This

[56] See, for example, *MTW* §21.2; Weinberg,[49] Chapter 12; J. L. Anderson, *Principles of Relativity Physics* (1967), §10-4.

is in marked contrast to the tradition established by Einstein himself, whose variational remarks are stylistically more similar to my own.[57]

But Lagrangian field theory is an elastic vessel. I sketch now an alternative approach due to Dirac.[58] Dirac proceeds from $\mathfrak{L} \sim \sqrt{g}\,R$ but (unlike Palatini) is prepared to assume at the outset that $\Gamma^{\sigma}{}_{\mu\nu} = \frac{1}{2}g^{\sigma\alpha}\{g_{\alpha\mu,\nu} + g_{\alpha\nu,\mu} - g_{\mu\nu,\alpha}\}$; i.e., that general relativity is an exercise in *Riemannian* geometry, nothing more abstruse. So he confronts a Lagrangian of the design $\mathfrak{L}(g, \partial g, \partial\partial g)$, but writing

$$R = \underbrace{g^{\mu\nu}(\Gamma^{\sigma}{}_{\mu\sigma,\nu} - \Gamma^{\sigma}{}_{\mu\nu,\sigma})}_{Q} - \underbrace{g^{\mu\nu}(\Gamma^{\sigma}{}_{\mu\nu}\Gamma^{\rho}{}_{\rho\sigma} - \Gamma^{\sigma}{}_{\mu\rho}\Gamma^{\rho}{}_{\nu\sigma})}_{R^*}$$

he observes that the offending $\partial\partial g$ terms are present only in the Q term (into which they enter linearly, and) from which they can be gauged away. More particularly, Dirac shows that

$$\sqrt{g}\,Q = 2\sqrt{g}\,R^* + \underbrace{(\sqrt{g}\,[g^{\rho\mu}\Gamma^{\sigma}{}_{\rho\sigma} - g^{\rho\sigma}\Gamma^{\mu}{}_{\rho\sigma}])_{,\mu}}_{\text{gauge term}}$$

so

$$\mathfrak{L} \sim \sqrt{g}\,R = \sqrt{g}\,R^* + \text{gauge term}$$

Dirac *abandons the gauge term* (where all $\partial\partial g$ terms reside), electing to work from

$$\begin{aligned}\mathfrak{L}^* \sim \sqrt{g}\,R^* = \tfrac{1}{4}\sqrt{g}\big\{&(g^{\mu\alpha}g^{\nu\beta} - g^{\mu\nu}g^{\alpha\beta})g^{\rho\sigma} \\ &- 2(g^{\mu\rho}g^{\alpha\beta} - g^{\mu\alpha}g^{\beta\rho})g^{\nu\sigma}\big\}g_{\mu\nu,\rho}g_{\alpha\beta,\sigma}\end{aligned} \tag{61}$$

which we see to be homogeneous of degree two in ∂g. This (because the rewards are so great) he is content to do even though $\mathfrak{L}^*$ *does not transform as a scalar density*; indeed, he considers the latter circumstance to be not a defect of the formalism but evidence that "four-dimensional symmetry is not a fundamental property of the physical world." Working from (61) he computes

$$\left\{\partial_\sigma \frac{\partial}{\partial g_{\mu\nu,\sigma}} - \frac{\partial}{\partial g_{\mu\nu}}\right\}\sqrt{g}\,R^* = -\sqrt{g}\,(R^{\mu\nu} - \tfrac{1}{2}Rg^{\mu\nu}) \tag{62.1}$$

---

[57] "Hamiltonsches Princip und allgemeine Relativitätstheorie" (1916), of which an English translation can be found in the Dover collection cited earlier.[52] Einstein cites, in addition to Hilbert, four papers by Lorentz (1915 & 1916).

[58] See Chapter 26—two and one half pages long—in his elegantly slim *General Theory of Relativity* (1996), to which I refer my reader for all the omitted details. Dirac presented his argument ("Theory of gravitation in Hamiltonian form," Proc. Roy. Soc. **A246**, 333 (1958)) for its methodological interest, as an illustrative application of ideas developed in "Generalized Hamiltonian dynamics," Proc. Roy. Soc. **A246**, 326 (1958).

I draw belated attention to the fact that in N-dimensional spacetime ($N \neq 2$)

$$R^{\mu\nu} - \tfrac{1}{2}Rg^{\mu\nu} = 0 \quad \Rightarrow \quad R = 0 \quad \Rightarrow \quad R^{\mu\nu} = 0 \tag{63}$$

With this fact in mind, Dirac observes that

$$\left\{\partial_\sigma \frac{\partial}{\partial \mathfrak{g}^{\mu\nu}{}_{,\sigma}} - \frac{\partial}{\partial \mathfrak{g}^{\mu\nu}}\right\}\sqrt{g}\, R^* = R_{\mu\nu} \tag{62.2}$$

Harmonic coordinates: Gravitational analog of the Lorentz gauge condition. By way of preparation, look back again to equations (2–24), where it is observed that the electromagnetic field equations can be expressed

$$\Box A^\nu - \partial^\nu(\partial_\mu A^\mu) = J^\nu \tag{64.1}$$

This looks like a system of four equations in four unknown fields A^μ. But the expressions on the left are (trivially and automatically) subject to a differential identity

$$\big[\Box A^\nu - \partial^\nu(\partial_\mu A^\mu)\big]_{,\nu} = 0 \tag{64.2}$$

—the solitary electromagnetic analog the contracted Bianchi identities (42.2). So the A^ν which satisfy (64.1) still *retain one degree of freedom*, familiar to us as gauge freedom $A^\nu \mapsto A'^\nu = A^\nu + \partial^\nu\chi$. That freedom can be exploited in various ways to achieve simplifications.[59] For example, we can install the Lorentz gauge condition $\partial_\nu A^\nu = 0$, replacing (64.1) by a *quintet* of equations

$$\Box A^\nu = J^\nu \quad \text{and} \quad \partial_\mu A^\mu = 0 \tag{64.3}$$

The identity (614.2) still pertains, but the A^ν which satisfy the expanded set of field equations are unique (retain no degrees of freedom).[60] Similarly...

The Einstein equations

$$R^{\mu\nu} - \tfrac{1}{2}Rg^{\mu\nu} = \kappa T^{\mu\nu} \tag{65.1}$$

[59] It might be interesting, on another occasion, to consider whether the non-Abelian gauge transformation

$$\boldsymbol{A}_\mu \longrightarrow \boldsymbol{A}'_\mu = \boldsymbol{S}\boldsymbol{A}_\mu\boldsymbol{S}^{-1} + i\,\tfrac{1}{g}\,\boldsymbol{S}_{,\mu}\boldsymbol{S}^{-1} \tag{3–93}$$

can be used to achieve useful simplification of

$$\left.\begin{aligned} \partial_\mu \boldsymbol{F}^{\mu\nu} &= \tfrac{1}{c}\boldsymbol{s}^\nu \\ \boldsymbol{s}^\nu &\equiv igc[\boldsymbol{F}^{\nu\alpha}, \boldsymbol{A}_\alpha] \end{aligned}\right\} \tag{3–106}$$

where $\boldsymbol{F}_{\mu\nu} \equiv (\partial_\mu \boldsymbol{A}_\nu - \partial_\nu \boldsymbol{A}_\mu) - ig(\boldsymbol{A}_\mu\boldsymbol{A}_\nu - \boldsymbol{A}_\nu\boldsymbol{A}_\mu)$.

[60] But see the cautionary statement two pages farther along!

appear on their face to be a system of ten equations in ten unknown fields $g^{\mu\nu}$. But the expressions on the left are (not quite trivially, but necessarily) subject to the quartet (42.2) of contracted Bianchi identities

$$\left[R^{\mu\nu} - \tfrac{1}{2}Rg^{\mu\nu}\right]_{;\,\nu} = 0 \tag{65.2}$$

So the $g^{\mu\nu}$ which satisfy (65.2) still *retain* $10 - 4 = 6$ *degrees of freedom*. This is associated in generally covariant theory with our freedom to recoordinatize, which is accomplished by the presentation of four (more or less) arbitrary functions: $x^\mu \mapsto \bar{x}^\mu = f^\mu(x)$.

To illustrate how "recoordinatization freedom" can be exploited to achieve simplifications, let φ be a scalar field and look to the generally covariant construction

$$\begin{aligned}\Box\varphi \equiv (g^{\mu\nu}\varphi_{;\,\nu})_{;\,\mu} &= (g^{\mu\nu}\varphi_{,\,\nu})_{;\,\mu} \\ &= (g^{\mu\nu}{}_{;\,\mu})\varphi_{,\nu} + g^{\mu\nu}(\varphi_{,\mu\nu} - \varphi_{,\alpha}\Gamma^\alpha{}_{\mu\nu})\end{aligned}$$

The argument which gave (59) can be tweaked to give $g^{\mu\nu}{}_{;\,\sigma} = 0$, so we have

$$= g^{\mu\nu}\varphi_{,\mu\nu} - \Gamma^\alpha\varphi_\alpha \tag{66.1}$$

$$\Gamma^\alpha \equiv g^{\mu\nu}\Gamma^\alpha{}_{\mu\nu} \tag{66.2}$$

It follows easily from (23) that Γ^α transforms by the rule

$$\bar{\Gamma}^\mu = \frac{\partial\bar{x}^\mu}{\partial x^\alpha}\Gamma^\alpha - g^{\alpha\beta}\frac{\partial^2\bar{x}^\mu}{\partial x^\alpha\partial x^\beta} \tag{66.3}$$

This last equation shows what (in principle) one must do to arrive in a coordinate system $\bar{x}$ in which $\bar{\Gamma}^\mu = 0$, and where, according to (66.1),

generally covariant d'Alembertian = ordinary d'Alembertian

One clearly *gives up general covariance to achieve such simplification*, but some freedom does survive: it follows from (66.3) that if $\Gamma^\alpha = 0$ then so does $\bar{\Gamma}^\mu = 0$ provided $x \mapsto \bar{x}$ is "harmonic" in the sense that

$$g^{\alpha\beta}\frac{\partial^2\bar{x}^\mu}{\partial x^\alpha\partial x^\beta} = 0$$

... All of which is precisely mimiced in electrodynamics, where in place of (66.3) one has

$$A'^\mu = A^\mu + \partial^\mu\chi$$

and to achieve the simplicity of Lorentz gauge $\partial_\mu A'^\mu = 0$ requires that χ be a solution of

$$\Box\chi = -\partial_\mu A^\mu$$

Gauge freedom has been sacrificed, but *some freedom does survive*: if $\partial_\mu A^\mu = 0$ then so does $\partial_\mu A'^\mu = 0$ provided $A^\mu \mapsto A'^\mu$ is "harmonic" in the sense that

$$\Box\chi = 0$$

How does this development square with the "uniqueness" claim made on the basis of (64.3)? It doesn't! ... the lesson being that

> *"degree of freedom counting" is delicate business, and generally unreliable unless side conditions (initial and boundary data) have been specified.*

Another way to characterize the simplification achieved by the introduction of harmonic coordinates: we have

$$\begin{aligned}
\Gamma^\sigma \equiv g^{\mu\nu}\Gamma^\sigma{}_{\mu\nu} &= \tfrac{1}{2}g^{\mu\nu}g^{\sigma\alpha}\left\{g_{\alpha\mu,\nu} + g_{\alpha\nu,\mu} - g_{\mu\nu,\alpha}\right\} \\
&= -\tfrac{1}{2}g^{\mu\nu}g_{\alpha\mu}g^{\sigma\alpha}{}_{,\nu} - \tfrac{1}{2}g^{\mu\nu}g_{\alpha\nu}g^{\sigma\alpha}{}_{,\mu} - \tfrac{1}{2}g^{\sigma\alpha}g^{\mu\nu}g_{\mu\nu,\alpha} \\
&= -g^{\sigma\alpha}{}_{,\alpha} - g^{\sigma\alpha}\cdot\tfrac{1}{\sqrt{g}}(\sqrt{g}\,)_{,\alpha} \quad \text{by (54)} \\
&= -\tfrac{1}{\sqrt{g}}(\sqrt{g}g^{\sigma\alpha})_{,\alpha}
\end{aligned}$$

From (56) follow the generally covariant statements

$$\mathfrak{g}^{\mu\alpha}{}_{;\alpha} = 0 \tag{67.1}$$

but the preceding manipulations show that in harmonic coordinate systems we have

$$\mathfrak{g}^{\mu\alpha}{}_{,\alpha} = 0 \tag{67.2}$$

These last conditions are trivially satisfied if $g^{\mu\nu}$ (whence also $\sqrt{g}$) are constant, as would be the case if, in the absence of gravitation, we installed Cartesian coordinates in flat space (or any coordinates harmonically related to them). We are not surprised, therefore, to find that harmonic coordinate systems lend themselves especially well to the description of *weak* gravitational fields, to discussion of the curved spacetime physics *in relation to flat spacetime physics*.[61]

It will be appreciated that, while the Principle of General Covariance served to guide the creation of general relativity, the abandonment of general covariance in favor of some specialized class of coordinate systems (harmonic coordinates, for example) *does no violence to the physics*: it has not the nature

[61] See, for example, *MTW*'s Chapter 18. It is curious that the index of the Black Bible contains no entry at "coordinates, harmonic;" what most authors refer to as "installation of harmonic coordinates" Misner, Thorne & Wheeler prefer to call "imposition of the Lorentz gauge condition"—ill-advisedly, in my view, for I think it deceptive to conflate *gauge freedom present in the physics written upon spacetime* with *freedom to recoordinatize spacetime itself*.

of an approximation... though it may facilitate the formulation of useful approximations. The situation here is similar to one encountered in classical mechanics, where imposition of a general covariance requirement leads from Newton's equations to Lagrange's, but the power of the Lagrangian formalism resides in the circumstance that it permits one to work in the special coordinates best suited to the specific problem at hand; it is in the spending that money reveals its value, in its abandonment that general covariance sometimes declares its utility.

5

CALCULUS OF FUNCTIONALS

Introduction. In classical mechanics one has interest in functions $x(t)$ of a single variable, while in field theory one's interest shifts to functions $\varphi(t,x)$ of several variables, but the ordinary calculus of functions of several variables would appear to be as adequate to the mathematical needs of the latter subject as it is to the former. And so, in large part, it is. But the "variational principles of mechanics" do typically reach a little beyond the bounds of the ordinary calculus. We had occasion already on p. 16 to remark in passing that $S[\varphi] = \int \mathcal{L}(\varphi, \partial\varphi)dx$ is by nature a "number-valued function of a function," and to speak therefore not of an "action function" but of the action *functional*; it was, in fact, to emphasize precisely that distinction that we wrote not $S(\varphi)$ but $S[\varphi]$. When one speaks—as Hamilton's principle asks us to do—of the variational properties of such an object one is venturing into the "calculus of functionals," but only a little way. Statements of the (frequently very useful) type $\delta f(x) = 0$ touch implicitly upon the concept of differentiation, but scarcely hint at the elaborate structure which is the calculus itself. Similarly, statements of the (also often useful) type $\delta S[\varphi] = 0$, though they exhaust the subject matter of the "calculus of variations," scarcely hint at the elaborate (if relatively little known) structure which I have in the present chapter undertaken to review. Concerning my motivation:

At (60) in Chapter I we had occasion to speak of the "Poisson bracket" of a pair of functionals $A = \int \mathcal{A}(\pi, \varphi, \nabla\pi, \nabla\varphi)dx$ and $B = \int \mathcal{B}(\pi, \varphi, \nabla\pi, \nabla\varphi)dx$, and this—since in mechanics the Poisson bracket of observables $A(p,q)$ and $B(p,q)$ is defined

$$[A,B] \equiv \sum_{k=1}^{n} \left\{ \frac{\partial A}{\partial q^k}\frac{\partial B}{\partial p_k} - \frac{\partial B}{\partial q^k}\frac{\partial A}{\partial p_k} \right\}$$

—would appear (as noted already on p. 49) to entail that be we in position to attach meaning to the "derivative" of a functional. One of my initial objectives

will be to show how this is done. We will learn how to compute functional derivatives to all orders, and will discover that such constructions draw in an essential way upon the theory of distributions—a subject which, in fact, belongs properly not to the theory of functions but to the theory of functionals.[1] We will find ourselves then in position to construct the functional analogs of Taylor series. Our practical interest in the integral aspects of the functional calculus we owe ultimately to Richard Feynman; we have seen that the Schrödinger equation provides—interpretive matters aside—a wonderful instance of a classical field, and it was Feynman who first noticed (or, at least, who first drew attention to the importance of the observation) that the Schrödinger equation describes a property of a certain functional integral. We will want to see how this comes about. Methods thus acquired become indispensible when one takes up the *quantization* of classical field theory; they are, in short, dominant in *quantum* field theory. Which, however, is not our present concern. Here our objective will be simply to put ourselves in position to recognize certain moderately subtle distinctions, to go about our field-theoretic business with greater precision, to express ourselves with greater notational economy.

Construction of the functional derivative. By way of preparation, let $F(\boldsymbol{\varphi})$ be a number-valued function of the finite-dimensional vector $\boldsymbol{\varphi}$. What shall we mean by the "derivative of $F(\boldsymbol{\varphi})$?" The key observation here is that we can mean *many* things, depending on the *direction in which we propose to move* while monitoring rate of change; the better question therefore is "What shall me mean by the *directional* derivative of $F(\boldsymbol{\varphi})$?" And here the answer is immediate: take $\boldsymbol{\lambda}$ to be any fixed vector and form

$$\begin{aligned} \mathbf{D}_{[\boldsymbol{\lambda}]}F(\boldsymbol{\varphi}) &\equiv \lim_{\epsilon\to 0}\frac{F(\boldsymbol{\varphi}+\epsilon\boldsymbol{\lambda})-F(\boldsymbol{\varphi})}{\epsilon} \qquad (1.1)\\ &= \sum_i \frac{\partial F(\boldsymbol{\varphi})}{\partial \varphi^i}\lambda^i \qquad (1.2)\\ &\equiv \text{derivative of } F(\boldsymbol{\varphi}) \text{ "in the direction } \boldsymbol{\lambda}\text{"} \end{aligned}$$

We might write

$$\mathbf{D}_{[\boldsymbol{\lambda}]}F(\boldsymbol{\varphi}) \equiv F[\boldsymbol{\varphi};\boldsymbol{\lambda}]$$

to underscore the fact that $\mathbf{D}_{[\boldsymbol{\lambda}]}F(\boldsymbol{\varphi})$ is a *bifunctional*. From (1.2) it is plain that $F[\boldsymbol{\varphi};\boldsymbol{\lambda}]$ is, so far as concerns its $\boldsymbol{\lambda}$-dependence, a *linear* functional:

$$F[\boldsymbol{\varphi};c_1\boldsymbol{\lambda}_1+c_2\boldsymbol{\lambda}_2] = c_1F[\boldsymbol{\varphi};\boldsymbol{\lambda}_1]+c_2F[\boldsymbol{\varphi};\boldsymbol{\lambda}_2]$$

I have now to describe the close relationship between "linear functionals" on the one hand, and "inner products" on the other. Suppose $\boldsymbol{\alpha}=\sum\alpha^i\mathbf{e}_i$ and $\boldsymbol{\lambda}=\sum\lambda^i\mathbf{e}_i$ are elements of an inner product space. Then $(\boldsymbol{\alpha},\boldsymbol{\lambda})$ is, by one

[1] See J. Lützen's wonderful little book, *The Prehistory of the Theory of Distributions*, (1982).

of the defining properties of the inner product, a bilinear functional. It is, in particular, a linear functional of $\boldsymbol{\lambda}$—call it $A[\boldsymbol{\lambda}]$—and can be described

$$(\boldsymbol{\alpha}, \boldsymbol{\lambda}) = A[\boldsymbol{\lambda}] = \sum_i \alpha_i \lambda^i \quad \text{with} \quad \alpha_i = \sum_j (\mathbf{e}_i, \mathbf{e}_j)\alpha^j$$

The so-called *Riesz-Frechet Theorem* runs in the opposite direction; it asserts[2] that if $A[\boldsymbol{\lambda}]$ is a linear functional defined on an inner product space, then there exists an $\boldsymbol{\alpha}$ such that $A[\boldsymbol{\lambda}]$ can be represented $A[\boldsymbol{\lambda}] = (\boldsymbol{\alpha}, \boldsymbol{\lambda})$. Returning in the light of this fundamental result to (1), we can say that (1.1) defines a linear functional of $\boldsymbol{\lambda}$, therefore there exists an $\boldsymbol{\alpha}$ such that

$$\mathbf{D}_{[\boldsymbol{\lambda}]} F(\boldsymbol{\varphi}) = (\boldsymbol{\alpha}, \boldsymbol{\lambda}) = \sum_i \alpha_i \lambda^i$$

and that we have found it convenient/natural in place of α_i to write $\frac{\partial F(\boldsymbol{\varphi})}{\partial \varphi^i}$.

It is by direct (infinite-dimensional) imitation of the preceeding line of argument that we construct what might be (but isn't) called the "directional derivative of a functional $F[\varphi(x)]$." We note that

$$\mathbf{D}_{[\lambda(x)]} F[\varphi(x)] \equiv \lim_{\epsilon \to 0} \frac{F[\varphi(x) + \epsilon\lambda(x)] - F[\varphi(x)]}{\epsilon} \tag{2.1}$$

is a linear functional of $\lambda(x)$, and admits therefore of the representation

$$= \int \alpha(x')\lambda(x')\, dx' \tag{2.2}$$

And we agree, as a matter simply of notation (more specifically, as a reminder that $\alpha(x)$ came into being as the result of a differentiation process applied to the functional $F[\varphi(x)]$), to write

$$\alpha(x) = \frac{\delta F[\varphi]}{\delta \varphi(x)} = \int \alpha(x')\, \delta(x' - x)\, dx' \tag{3}$$

Evidently $\frac{\delta F[\varphi]}{\delta \varphi(x)}$ itself can be construed to describe the result of differentiating $F[\varphi]$ in the "direction" of the δ-function which is singular at the point x. If, in particular, $F[\varphi]$ has the structure

$$F[\varphi] = \int f(\varphi(x))\, dx \tag{4.1}$$

then the construction (2.1) gives

$$\mathbf{D}_{[\lambda]} F[\varphi] = \int \lambda(x) \frac{\partial}{\partial \varphi} f(\varphi(x))\, dx$$

[2] For a proof see F. Riesz & B. Sz.Nagy, *Functional Analysis* (1955), p. 61.

whence

$$\frac{\delta F[\varphi]}{\delta\varphi(x)} = \frac{\partial}{\partial\varphi} f(\varphi(x)) \tag{4.2}$$

And if, more generally,

$$F[\varphi] = \int f(\varphi(x), \varphi_x(x))\, dx \tag{5.1}$$

then by a familiar argument

$$\mathbf{D}_{[\lambda]} F[\varphi] = \int \lambda(x) \left\{ \frac{\partial}{\partial\varphi} - \frac{\partial}{\partial x}\frac{\partial}{\partial\varphi_x} \right\} f(\varphi(x), \varphi_x(x))\, dx$$

and we have

$$\frac{\delta F[\varphi]}{\delta\varphi(x)} = \left\{ \frac{\partial}{\partial\varphi} - \frac{\partial}{\partial x}\frac{\partial}{\partial\varphi_x} \right\} f(\varphi(x), \varphi_x(x)) \tag{5.2}$$

In such cases $\frac{\delta F[\varphi]}{\delta\varphi(x)}$ flourishes in the sunshine, as it has flourished unrecognized in preceeding pages, beginning at about p. 43. But it is in the general case distribution-like, as will presently become quite clear; it popped up, at (2.2), in the shade of an integral sign, and can like a mushroom become dangerous when removed from that protective gloom. There is, in short, some hazard latent in the too-casual use of (3).

It follows readily from (2) that the functional differentiation operator $\mathbf{D}_{[\lambda(x)]}$ acts linearly

$$\mathbf{D}_{[\lambda]}\big\{F[\varphi] + G[\varphi]\big\} = \mathbf{D}_{[\lambda]} F[\varphi] + \mathbf{D}_{[\lambda]} G[\varphi] \tag{6.1}$$

and acts on products by the familiar rule

$$\mathbf{D}_{[\lambda]}\big\{F[\varphi]\cdot G[\varphi]\big\} = \big\{\mathbf{D}_{[\lambda]} F[\varphi]\big\}\cdot G[\varphi] + F[\varphi]\cdot\big\{\mathbf{D}_{[\lambda]} G[\varphi]\big\} \tag{6.2}$$

In the shade of an implied integral sign we therefore have

$$\frac{\delta}{\delta\varphi(x)}\Big\{F[\varphi] + G[\varphi]\Big\} = \frac{\delta}{\delta\varphi(x)} F[\varphi] + \frac{\delta}{\delta\varphi(x)} G[\varphi] \tag{7.1}$$

and

$$\frac{\delta}{\delta\varphi(x)}\Big\{F[\varphi]\cdot G[\varphi]\Big\} = \Big\{\frac{\delta}{\delta\varphi(x)} F[\varphi]\Big\}\cdot G[\varphi] + F[\varphi]\cdot\Big\{\frac{\delta}{\delta\varphi(x)} G[\varphi]\Big\} \tag{7.2}$$

In connection with the product rule it is, however, important not to confuse $F[\varphi]\cdot G[\varphi] = \int f dx' \cdot \int g\, dx''$ with $F[\varphi] * G[\varphi] = \int (f\cdot g)\, dx$.

A second φ-differentiation of the linear bifunctional

$$F[\varphi;\lambda_1] = \mathbf{D}_{[\lambda_1]} F[\varphi] = \int \frac{\delta F[\varphi]}{\delta\varphi(x')}\lambda_1(x')\, dx'$$

yields a bilinear trifunctional

$$F[\varphi;\lambda_1,\lambda_2] = \mathbf{D}_{[\lambda_2]}F[\varphi;\lambda_1] = \mathbf{D}_{[\lambda_2]}\mathbf{D}_{[\lambda_1]}F[\varphi] \\ = \int\int \frac{\delta^2 F[\varphi]}{\delta\varphi(x')\delta\varphi(x'')}\lambda_1(x')\lambda_2(x'')\,dx'dx'' \tag{8}$$

In general we expect to have (and will explicitly assume that)

$$\mathbf{D}_{[\lambda_2]}\mathbf{D}_{[\lambda_1]}F[\varphi] = \mathbf{D}_{[\lambda_1]}\mathbf{D}_{[\lambda_2]}F[\varphi]$$

which entails that

$$\frac{\delta^2 F[\varphi]}{\delta\varphi(x')\delta\varphi(x'')} \quad \text{is a } \textit{symmetric} \text{ function of } x' \text{ and } x''$$

By natural extension we construct φ-derivatives of all orders:

$$\mathbf{D}_{[\lambda_n]}\cdots\mathbf{D}_{[\lambda_2]}\mathbf{D}_{[\lambda_1]}F[\varphi] = F[\varphi;\lambda_1,\lambda_2,\ldots,\lambda_n] \\ = \iint\cdots\int \underbrace{\frac{\delta^n F[\varphi]}{\delta\varphi(x^1)\delta\varphi(x^2)\cdots\delta\varphi(x^n)}}_{\text{totally symmetric in } x^1,x^2,\ldots,x^n} \lambda_1(x^1)\lambda_2(x^2)\cdots\lambda_2(x^2)\,dx^1dx^2\cdots dx^n$$

Functional analog of Taylor's series. In the ordinary calculus of functions, derivatives of ascending order are most commonly encountered in connection with the theory of Taylor series; one writes

$$f(x+a) = \exp\Big\{a\frac{d}{dx}\Big\}f(x) = \sum_n \frac{1}{n!}\frac{d^n f(x)}{dx^n}a^n \tag{9}$$

which is justified by the observation that, for all n,

$$\lim_{a\to 0}\Big(\frac{d}{da}\Big)^n(\text{lefthand side}) = \lim_{a\to 0}\Big(\frac{d}{da}\Big)^n(\text{righthand side})$$

Similarly

$$f(x+a,y+b) = \exp\Big\{a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y}\Big\}f(x) \\ = f(x,y) + \big\{af_x(x,y) + bf_y(x,y)\big\} \\ + \tfrac{1}{2!}\big\{a^2 f_{xx}(x,y) + 2abf_{xy}(x,y) + b^2 f_{yy}(x,y)\big\} + \cdots$$

No mystery attaches now to the sense in which (and why) it becomes possible (if we revert to the notation of p.74) to write

$$F(\boldsymbol{\varphi}+\boldsymbol{\lambda}) = \exp\Big\{\sum \lambda^i \frac{\partial}{\partial\varphi^i}\Big\}F(\boldsymbol{\varphi}) \\ = \sum_{n=0}^{\infty}\frac{1}{n!}\Big\{\sum_{i_1}\sum_{i_2}\cdots\sum_{i_n} F_{i_1 i_2\cdots i_n}(\boldsymbol{\varphi})\lambda^{i_1}\lambda^{i_2}\cdots\lambda^{i_n}\Big\}$$

with $F_{i_1 i_2 \cdots i_n}(\boldsymbol{\varphi}) \equiv \partial_{i_1}\partial_{i_2}\cdots\partial_{i_n}F(\boldsymbol{\varphi})$ where $\partial_i \equiv \frac{\partial}{\partial\varphi^i}$. Passing finally to the continuous limit, we obtain

$$F[\varphi+\lambda] = \exp\left\{\mathbf{D}_{[\lambda]}\right\}F[\varphi] \tag{10}$$
$$= \sum_{n=0}^{\infty}\frac{1}{n!}\left\{\iint\cdots\int F(x^1,x^2,\ldots,x^n;\varphi)\lambda(x^1)\lambda(x^2)\cdots\lambda(x^n)\,dx^1dx^2\cdots dx^n\right\}$$

with $F(x^1,x^2,\ldots,x^n;\varphi) \equiv \delta^n F[\varphi]/\delta\varphi(x^1)\delta\varphi(x^2)\cdots\delta\varphi(x^n)$. The right side of (10) displays $F[\varphi+\lambda]$ as a "Volterra series"—the functional counterpart of a Taylor series.[3] Taylor's formula (11) embodies an idea

$$\text{function of interest} = \sum \text{elementary functions}$$

which is well known to lie at the heart of analysis (i.e., of the theory of functions).[4] We are encouraged by the abstractly identical structure and intent of Volterra's formula (12) to hope that this obvious variant

$$\text{functional of interest} = \sum \text{elementary functionals}$$

[3] Set $a = \delta x$ in (9), or $\lambda = \delta\varphi$ in (10), and you will appreciate that I had the calculus of variations in the back of my mind when I wrote out the preceeding material. The results achieved seem, however, "backwards" from another point of view; we standardly seek to "expand $f(a+x)$ in powers of x about the point a," not the reverse. A version of (9) which is less offensive to the eye of an "expansion theorist" can be achieved by simple interchange $x \leftrightharpoons a$:

$$f(a+x) = \sum_n \frac{1}{n!}f_n x^n \quad \text{with} \quad f_n = \frac{d^n f(a)}{da^n} \tag{11}$$

Similarly, (10) upon interchange $\varphi \leftrightharpoons \lambda$ and subsequent notational adjustment $\lambda \to \alpha$ becomes

$$F[\alpha+\varphi] = \sum_{n=0}^{\infty}\frac{1}{n!}\left\{\iint\cdots\int F(x^1,x^2,\ldots,x^n)\cdot\right. \tag{12}$$
$$\left.\cdot\,\varphi(x^1)\varphi(x^2)\cdots\varphi(x^n)\,dx^1dx^2\cdots dx^n\right\}$$

with $F(x^1,x^2,\ldots,x^n) \equiv \delta^n F[\alpha]/\delta\alpha(x^1)\delta\alpha(x^2)\cdots\delta\alpha(x^n)$.

[4] To Kronecker's "God made the integers; all else is the work of man" his contemporary Weierstrass is said to have rejoined "God made *power series*; all else is the work of man." It becomes interesting in this connection to recall that E. T. Bell, in his *Men of Mathematics*, introduces his Weierstrass chapter with these words: "The theory that has had the greatest development in recent times is without any doubt the theory of functions." And that those words were written by Vito Volterra—father of the functional calculus.

of that same idea will lead with similar force and efficiency to a theory of functionals (or functional analysis).

Simple examples serve to alert us the fact that surprises await the person who adventures down such a path. Look, by way of illustration, to the functional

$$F[\varphi] \equiv \int \varphi^2(x)\,dx \tag{13.1}$$

Evidently we can, if we wish, write

$$= \int\int F(x^1, x^2) \cdot \varphi(x^1)\varphi(x^2)\,dx^1 dx^2$$

with

$$F(x^1, x^2) = \delta(x^1 - x^2) \tag{13.2}$$

The obvious lesson here is that, while the expansion coefficients f_n which appear on the right side of Taylor's formula (11) are by nature numbers, the "expansion coefficients" $F(x^1, x^2, \ldots, x^n)$ which enter into the construction (12) of a "Volterra series" may, even in simple cases, be *distributions*.

Look next to the example

$$F[\varphi] \equiv \int \varphi(x)\varphi_x(x)\,dx \tag{14.1}$$

From $\varphi\varphi_x = \frac{1}{2}(\varphi^2)_x$ we obtain

$$= \text{boundary terms}$$

Since familiar hypotheses are sufficient to cause boundary terms to vanish, what we have here is a *complicated description of the zero-functional*, the Volterra expansion of which is trivial. One might, alternatively, argue as follows:

$$\begin{aligned}\varphi_x(x) &= \int \delta(y-x)\varphi_y(y)\,dy \\ &= -\int \delta'(y-x)\varphi(y)\,dy \quad \text{by partial integration}\end{aligned}$$

Therefore

$$F[\varphi] = -\int\int \delta'(y-x)\varphi(x)\varphi(y)\,dxdy$$

But only the symmetric part of $\delta'(y-x)$ contributes to the double integral, and

$$\delta'(y-x) = \lim_{\epsilon\to 0} \frac{\delta(y-x+\epsilon) - \delta(y-x-\epsilon)}{2\epsilon}$$

is, by the symmetry of $\delta(z)$, clearly antisymmetric. So again we obtain

$$F[\varphi] = 0 \tag{14.2}$$

Look finally to the example

$$F[\varphi] \equiv \int \varphi_x^2(x)\,dx \tag{15.1}$$

From $\varphi_x^2 = -\varphi\varphi_{xx} + (\varphi\varphi_x)_x$ we have $F[\varphi] = -\int \varphi\varphi_{xx}\,dx$ plus a boundary term which we discard. But

$$\begin{aligned}\varphi_{xx}(x) &= \int \delta(y-x)\varphi_{yy}(y)\,dy \\ &= -\int \delta'(y-x)\varphi_y(y)\,dy \quad \text{by partial integration} \\ &= +\int \delta''(y-x)\varphi(y)\,dy \quad \text{by a second partial integration}\end{aligned}$$

and

$$\delta''(y-x) = \lim_{\epsilon\to 0} \frac{\delta'(y-x+\epsilon) - \delta'(y-x-\epsilon)}{2\epsilon}$$

is, by the antisymmetry of $\delta'(z)$, symmetric. So we have

$$F[\varphi] = \int\int F(x^1, x^2)\cdot\varphi(x^1)\varphi(x^2)\,dx^1 dx^2$$

with

$$F(x^1, x^2) = \delta''(x^1 - x^2) \tag{15.2}$$

Again, the "expansion coefficient" $F(x^1, x^2)$ is not a number, not a function, but a distribution.

The methods employed in the treatment of the preceeding examples have been straightforward enough, yet somewhat *ad hoc*. Volterra's formula (12) purports to supply a systematic general method for attacking such problems. How does it work? Let $F[\varphi]$ be assumed to have the specialized but frequently encountered form

$$F[\varphi] = \int f(\varphi, \varphi_x)\,dx \tag{16}$$

Our objective is to expand $F[\alpha+\varphi]$ "in powers of φ" (so to speak) and then set $\alpha(x)$ equal to the "zero function" $0(x)$. Our objective, in short, is to construct the functional analog of a "Maclaurin series:"

$$F[\varphi] = F_0 + F_1[\varphi] + \tfrac{1}{2}F_2[\varphi] + \cdots \tag{17}$$

Trivially $F_0 = \int f(\alpha, \alpha_x)\,dx\big|_{\alpha=0}$ and familiarly

$$\begin{aligned}F_1[\varphi] &= \left[\mathbf{D}_{[\varphi]} \int f(\alpha,\alpha_x)\,dx\right]_{\alpha=0} \\ &= \int \underbrace{\left[\left\{\frac{\partial}{\partial\alpha} - \frac{d}{dx}\frac{\partial}{\partial\alpha_x}\right\} f(\alpha,\alpha_x)\right]_{\alpha=0}}_{= \dfrac{\delta F[\alpha]}{\delta\alpha(x)},\ \text{evaluated at } \alpha=0} \varphi(x)\,dx\end{aligned} \tag{18}$$

But this is as far as we can go on the basis of experience standard to the (1^{st}-order) calculus of variations. We have

$$\begin{aligned} F_2[\varphi] &= \left[\mathbf{D}_{[\varphi]}\mathbf{D}_{[\varphi]}\int f(\alpha,\alpha_x)\,dx\right]_{\alpha=0} \\ &= \left[\mathbf{D}_{[\varphi]}\int \underbrace{g(\alpha,\alpha_x,\alpha_{xx})}\varphi(x)\,dx\right]_{\alpha=0} \qquad (19) \\ &\qquad\qquad = \left\{\frac{\partial}{\partial\alpha}-\frac{d}{dx}\frac{\partial}{\partial\alpha_x}\right\}f(\alpha,\alpha_x) = \frac{\delta F[\alpha]}{\delta\alpha(x)} \\ &= \left[\int \frac{\delta^2 F[\alpha]}{\delta\alpha(x)\delta\alpha(y)}\varphi(x)\varphi(y)\,dxdy)\right]_{\alpha=0} \end{aligned}$$

Our problem is construct explicit descriptions of the 2^{nd} variational derivative $\frac{\delta^2 F[\alpha]}{\delta\alpha(x)\delta\alpha(y)}$ and of its higher-order counterparts. The trick here is to introduce

$$\varphi(x) = \int \delta(x-y)\varphi(y)\,dx \qquad (20)$$

into (19) for we then have

$$\begin{aligned} F_2[\varphi] = \int \underbrace{\left[\mathbf{D}_{[\varphi]}\int g(\alpha,\alpha_x,\alpha_{xx})\delta(x-y)\,dx\right]_{\alpha=0}}\varphi(y)\,dy \\ = \left[\int h(\alpha,\alpha_x,\alpha_{xx},\alpha_{xxx})\,\varphi(x)\,dx\right]_{\alpha=0} \qquad (21.1) \end{aligned}$$

with

$$\begin{aligned} h(\alpha,\alpha_x,\alpha_{xx},\alpha_{xxx}) &= \frac{\delta^2 F[\alpha]}{\delta\alpha(x)\delta\alpha(y)} \qquad (21.2) \\ &= \left\{\frac{\partial}{\partial\alpha}-\frac{d}{dx}\frac{\partial}{\partial\alpha_x}+\frac{d^2}{dx^2}\frac{\partial}{\partial\alpha_{xx}}\right\}g(\alpha,\alpha_x,\alpha_{xx})\delta(x-y) \end{aligned}$$

The pattern of the argument should at this point be clear; as one ascends from one order to the next one begins always by invoking (20), with the result that δ-functions and their derivatives stack ever deeper, while the variational derivative operator

$$\left\{\frac{\partial}{\partial\alpha}-\frac{d}{dx}\frac{\partial}{\partial\alpha_x}+\frac{d^2}{dx^2}\frac{\partial}{\partial\alpha_{xx}}+\cdots\right\}$$

acquires at each step one additional term.

In ordinary analysis one can expand $f(x)$ about the point $x=a$ only if a is a "regular point," and the resulting Taylor series $f(x)=\sum\frac{1}{n!}f_n(x-a)^n$ can be expected to make sense only within a certain "radius of convergence." Such details become most transparent when x is allowed to range on the complex

plane. Similar issues—though I certainly do not intend to pursue them here—can be expected to figure prominently in any fully developed account of the theory of functionals.

Construction of functional analogs of the Laplacian and Poisson bracket. Now that we possess the rudiments of a theory of functional differentiation, we are in position to contemplate a "theory of functional differential equations." I do not propose to lead an expedition into that vast jungle, which remains (so far as I am aware) still largely unexplored. But I do want to step out of our canoe long enough to draw your attention to one small methodological flower that grows there on the river bank, at the very edge of the jungle. Central to many of the partial differential equations of physics is the Laplacian operator, ∇^2. Here in the jungle it possess a functional analog. How is such an object to be constructed? The answer is latent in the "sophisticated" answer to a simpler question, which I pose in the notation of p. 74: How does

$$\nabla^2 F(\boldsymbol{\varphi}) = \left\{ \sum_i \left(\frac{\partial}{\partial \varphi^i} \right)^2 \right\} F = \mathrm{tr} \| \partial^2 F / \partial \varphi^i \partial \varphi^j \|$$

come to acquire from $\mathbf{D}_{[\boldsymbol{\lambda}]} F(\boldsymbol{\varphi})$ the status of a "natural object"? Why, in particular, does $\nabla^2 F$ contain no allusion to $\boldsymbol{\lambda}$? We proceed from the observation that

$$\begin{aligned} \mathbf{D}_{[\boldsymbol{\lambda}]} \mathbf{D}_{[\boldsymbol{\lambda}]} F(\boldsymbol{\varphi}) &= \sum \sum \frac{\partial^2 F}{\partial \varphi^i \partial \varphi^j} \lambda^i \lambda^j \\ &= \boldsymbol{\lambda}^{\mathsf{T}} \mathbb{F} \, \boldsymbol{\lambda} \qquad \text{where } \mathbb{F} \text{ is the square matrix } \| \partial^2 F / \partial \varphi^i \partial \varphi^j \| \\ &= \mathrm{tr} \, \mathbb{F} \mathbb{L} \qquad \text{where } \mathbb{L} \text{ is the square matrix } \| \lambda^i \lambda^j \| \end{aligned}$$

We note more particularly that $\mathbb{L}^2 = (\boldsymbol{\lambda} \cdot \boldsymbol{\lambda}) \mathbb{L}$, so if $\boldsymbol{\lambda}$ is a unit vector ($\boldsymbol{\lambda} \cdot \boldsymbol{\lambda} = 1$) then $\mathbb{L}$ is a *projection matrix* which in fact projects upon $\boldsymbol{\lambda}$: $\mathbb{L}\mathbf{x} = (\boldsymbol{\lambda} \cdot \mathbf{x}) \boldsymbol{\lambda}$. Now let $\{\mathbf{e}_i\}$ refer to some (any) *orthonormal basis*, and let $\{\mathbb{E}_i\}$ denote the associated set of projection matrices. Orthonormality entails $\mathbb{E}_i \mathbb{E}_j = \delta_{ij} \mathbb{E}_i$ while $\sum \mathbb{E}_i = \mathbb{I}$ expresses the presumed *completeness* of the set $\{\mathbf{e}_i\}$. From these elementary observations[5] it follows that

$$\begin{aligned} \sum \mathbf{D}_{[\mathbf{e}_i]} \mathbf{D}_{[\mathbf{e}_i]} F(\boldsymbol{\varphi}) = \sum \mathrm{tr} \, \mathbb{F} \mathbb{E}_i &= \mathrm{tr} \, \mathbb{F} \\ &= \sum \| \partial^2 F / \partial \varphi^i \partial \varphi^i \| \\ &= \nabla^2 F(\boldsymbol{\varphi}) \end{aligned}$$

Turning now from functions to functionals, we find it quite natural to construct

$$\sum \mathbf{D}_{[e_i]} \mathbf{D}_{[e_i]} F[\varphi] = \sum \int \int \frac{\delta^2 F[\varphi]}{\delta \varphi(x) \delta \varphi(y)} e_i(x) e_i(y) \, dx dy$$

[5] The ideas assembled here acquire a striking transparency when formulated in a simplified variant of Dirac's "bra-ket notation." Readers familiar with that notation are encouraged to give it a try.

And if we assume the functions $\{e_i(x)\}$ to be orthonormal $\int e_i(x)e_j(x)\,dx = \delta_{ij}$ and (which is more to the immediate point) complete in function space

$$\sum e_i(x)e_j(y) = \delta(x-y)$$

then we obtain this natural definition of the "Laplacian of a functional":

$$\boldsymbol{\nabla}^2 F[\varphi] \equiv \Big\{ \sum \mathbf{D}_{[e_i]}\mathbf{D}_{[e_i]} \Big\} F[\varphi] = \int\int \frac{\delta^2 F[\varphi]}{\delta\varphi(x)\delta\varphi(x)}\,dx \tag{22}$$

Familiarly, the "partial derivative of a function of several variables" is a concept which arises by straightforward generalization from that of "the (ordinary) derivative of a function of a single variable." The "partial functional derivative" springs with similar naturalness from the theory of "ordinary functional derivatives," as outlined in preceeding paragraphs; the problem one encounters is not conceptual but notational/terminological. Let us agree to write $\mathbf{D}_{[\lambda]/\varphi}F[\ldots,\varphi,\ldots]$ to signify "the partial derivative of $F[\ldots,\varphi,\ldots]$—a functional of several variables—*with respect to* $\varphi(x)$ *in the direction* $\lambda(x)$:

$$\mathbf{D}_{[\lambda]/\varphi}F[\ldots,\varphi,\ldots] = \int \frac{\delta F[\ldots,\varphi,\ldots]}{\delta\varphi(x)}\lambda(x)\,dx$$

I shall not pursue this topic to its tedious conclusion, save to remark that one expects quite generally to have "equality of cross derivatives"

$$\mathbf{D}_{[\lambda_1]/\varphi}\mathbf{D}_{[\lambda_2]/\psi}F[\varphi,\psi] = \mathbf{D}_{[\lambda_2]/\psi}\mathbf{D}_{[\lambda_1]/\varphi}F[\varphi,\psi]$$

since, whether one works from the expression on the left or from that on the right, one encounters $F[\varphi+\lambda_1,\psi+\lambda_2] - F[\varphi+\lambda_1,\psi] - F[\varphi,\psi+\lambda_2] + F[\varphi,\psi]$. Instead I look to one of the problems that, on p. 73, served ostensibly to motivate this entire discussion. Let $A[\varphi,\pi]$ and $B[\varphi,\pi]$ be given functionals of two variables, and construct

$$\begin{aligned}\mathbf{D}_{[\lambda_1]/\varphi}A\cdot\mathbf{D}_{[\lambda_2]/\pi}B - \mathbf{D}_{[\lambda_1]/\varphi}B\cdot\mathbf{D}_{[\lambda_2]/\pi}A& \\ = \int\int \left\{ \frac{\delta A}{\delta\varphi(x)}\frac{\delta B}{\delta\pi(y)} - \frac{\delta B}{\delta\varphi(x)}\frac{\delta A}{\delta\pi(y)} \right\}\lambda_1(x)\lambda_2(y)\,dxdy&\end{aligned}$$

Proceeding now in direct imitation of the construction which led us a moment ago to the definition (22) of the functional Laplacian, we write

$$\begin{aligned}[A,B] &\equiv \sum \left\{ \mathbf{D}_{[e_i]/\varphi}A\cdot\mathbf{D}_{[e_i]/\pi}B - \mathbf{D}_{[e_i]/\varphi}B\cdot\mathbf{D}_{[e_i]/\pi}A \right\} \\ &= \int\int \left\{ \frac{\delta A}{\delta\varphi(x)}\frac{\delta B}{\delta\pi(y)} - \frac{\delta B}{\delta\varphi(x)}\frac{\delta A}{\delta\pi(y)} \right\}\cdot\sum e_i(x)e_i(y)\,dxdy \\ &= \int \left\{ \frac{\delta A}{\delta\varphi(x)}\frac{\delta B}{\delta\pi(x)} - \frac{\delta B}{\delta\varphi(x)}\frac{\delta A}{\delta\pi(x)} \right\}dx \\ &= \int [\mathcal{A},\mathcal{B}]\,dx \quad \text{in the notation of (60), Chapter I}\end{aligned} \tag{23}$$

At (23) we have achieved our goal; we have shown that the Poisson bracket —a construct fundamental to Hamiltonian field theory—admits of natural description in language supplied by the differential calculus of functionals.

I return at this point to discussion of our "functional Laplacian," partly to develop some results of intrinsic interest, and partly to prepare for subsequent discussion of the integral calculus of functionals. It is a familiar fact—the upshot of a "folk theorem"—that

$$\nabla^2\varphi(x) \sim \langle\text{average of neighboring values}\rangle - \varphi(x) \tag{24}$$

$\nabla^2\varphi$ provides, therefore, a natural measure of the "local disequilibration" of the φ-field; in the absence of "disequilibration" ($\nabla^2\varphi = 0$) the field is said to be "harmonic." I begin with discussion of an approach to the proof of (24) which, though doubtless "well-known" in some circles, occurred to me only recently.[6] Working initially in two dimensions, let $\varphi(x, y)$ be defined on a neighborhood containing the point (x, y) on the Euclidian plane. At points on the boundary of a disk centered at (x, y) the value of φ is given by

$$\begin{aligned}\varphi(x + r\cos\theta, y + r\sin\theta) &= e^{r\cos\theta\frac{\partial}{\partial x} + r\sin\theta\frac{\partial}{\partial y}} \cdot \varphi(x, y)\\ &= \varphi + r(\varphi_x\cos\theta + \varphi_y\sin\theta)\\ &\quad + \tfrac{1}{2}r^2(\varphi_{xx}\cos^2\theta + 2\varphi_{xy}\cos\theta\sin\theta + \varphi_{yy}\sin^2\theta) + \cdots\end{aligned}$$

The average $\langle\varphi\rangle$ of the values assumed by φ on the boundary of the disk is given therefore by

$$\begin{aligned}\langle\varphi\rangle &= \frac{1}{2\pi r}\int_0^{2\pi}\{\text{right side of preceeding equation}\}\,r d\theta\\ &= \varphi + 0 + \tfrac{1}{4}r^2\{\varphi_{xx} + \varphi_{yy}\} + \cdots\end{aligned}$$

from which we obtain

$$\nabla^2\varphi = \tfrac{4}{r^2}\big\{\langle\varphi\rangle - \varphi\,\big\} + \cdots \quad \text{in the 2-dimensional case} \tag{25}$$

This can be read as a sharpened instance of (24).[7] In three dimensions we are motivated to pay closer attention to the notational organization of the

[6] See the introduction to my notes "Applications of the Theory of Harmonic Polynomials" for the Reed College Math Seminar of 7 March 1996.

[7] If φ refers physically to (say) the displacement of a membrane, then it becomes natural to set

$$\begin{aligned}\text{restoring force} &= k\big\{\langle\varphi\rangle - \varphi\,\big\}\\ &= \text{mass element}\cdot\text{acceleration}\\ &= 2\pi r^2\rho\cdot\varphi_{tt}\end{aligned}$$

and we are led from (25) to an instance of the wave equation:

$$\nabla^2\varphi = \tfrac{1}{c^2}\varphi_{tt}$$

with $c^2 = k/8\pi\rho$.

argument; we write

$$\varphi(\boldsymbol{x}+\boldsymbol{r}) = \varphi(\boldsymbol{x}) + \boldsymbol{r}\cdot\boldsymbol{\nabla}\varphi + \tfrac{1}{2}\boldsymbol{r}\cdot\begin{pmatrix}\varphi_{11} & \varphi_{12} & \varphi_{13}\\ \varphi_{21} & \varphi_{22} & \varphi_{23}\\ \varphi_{31} & \varphi_{32} & \varphi_{33}\end{pmatrix}\boldsymbol{r} + \cdots \tag{26}$$

which we want to average over the surface of the sphere $r_1^2 + r_2^2 + r_3^2 = r^2$. It is to that end that I digress now to establish a little LEMMA, which I will phrase with an eye to its dimensional generalization:

Let $\langle x^p\rangle$ denote the result of averaging the values assumed by x^p on the surface of the 3-sphere of radius r. Proceeding in reference to the figure, we have

$$\langle x^p\rangle = \frac{1}{S_3(r)}\int x^p\, dS$$

$$dS = S_2(r\sin\theta)\cdot r\,d\theta$$

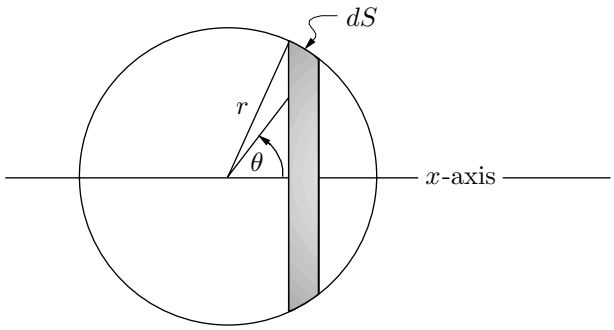


FIGURE 1: *Geometrical construction intended to set notation used in proof of the Lemma. The figure pertains fairly literally to the problem of averaging x^p over the surface of a 3-sphere, but serves also to provide schematic illustration of our approach to the problem of averaging x^p over the surface of an N-sphere.*

where I have adopted the notation

$$S_N(r) \equiv \text{surface area of } N\text{-sphere of radius } r = \begin{cases} 2\pi r & \text{when } N=2\\ 4\pi r^2 & \text{when } N=3\\ \vdots & \end{cases}$$

Evidently

$$\langle x^p\rangle = \frac{S_2(r)}{S_3(r)}\, r^{p+1}\underbrace{\int_0^\pi \cos^p\theta\,\sin\theta\, d\theta}$$

$$= \int_{-1}^{1} u^p\, du = \begin{cases} \frac{2}{p+1} & \text{for } p \text{ even}\\ 0 & \text{for } p \text{ odd}\end{cases}$$

which in the cases of particular interest gives

$$\begin{aligned}\langle x^0\rangle &= \frac{S_2(r)}{S_3(r)}\,r^1\,\frac{2}{1} = 1\\ \langle x\ \rangle &= 0\\ \langle x^2\rangle &= \frac{S_2(r)}{S_3(r)}\,r^3\,\frac{2}{3} = \tfrac{1}{3}r^2\end{aligned}$$

Returning with this information to (26), we rotate to the coordinate system relative to which the $\|\varphi_{ij}\|$ matrix is diagonal

$$\begin{pmatrix}\varphi_{11} & \varphi_{12} & \varphi_{13}\\ \varphi_{21} & \varphi_{22} & \varphi_{23}\\ \varphi_{31} & \varphi_{32} & \varphi_{33}\end{pmatrix} \longrightarrow \begin{pmatrix}\phi_{11} & 0 & 0\\ 0 & \phi_{22} & 0\\ 0 & 0 & \phi_{33}\end{pmatrix}$$

and obtain

$$\langle\varphi\rangle = \varphi + 0 + \tfrac{1}{2}\mathrm{tr}\begin{pmatrix}\phi_{11} & 0 & 0\\ 0 & \phi_{22} & 0\\ 0 & 0 & \phi_{33}\end{pmatrix}\cdot\tfrac{1}{3}r^2$$

But the trace is rotationally invariant, so we have (compare (25))

$$\nabla^2\varphi = \tfrac{6}{r^2}\big\{\langle\varphi\rangle - \varphi\ \big\} + \cdots \quad \text{in the 3-dimensional case} \tag{27}$$

Dimensional generalization of this result follows trivially upon dimensional generalization of our LEMMA. If $\langle x^p\rangle$ is taken now to denote the result of averaging the values assumed by x^p on the surface of the N-sphere of radius r, then—arguing as before—we have

$$\langle x^p\rangle = \frac{1}{S_N(r)}\int x^p\,dS$$

$$dS = S_{N-1}(r\sin\theta)\cdot r\,d\theta$$

A simple scaling argument is sufficient to establish that

$$S_N(r) = r^{N-1}\cdot S_N(1)$$

so we have

$$\langle x^p\rangle = \frac{S_{N-1}(1)}{S_N(1)}\,r^p\underbrace{\int_0^\pi \cos^p\theta\,\sin^{N-2}\theta\,d\theta} \tag{28}$$

and because the integrand is even/odd on the interval $0\le\theta\le\pi$ we have (for $N = 2,3,4,\ldots$; i.e., for $q\equiv N-2 = 0,1,2,\ldots$)

$$= \begin{cases} 0 & \text{when } p \text{ is odd}\\ 2\displaystyle\int_0^{\frac{1}{2}\pi}\cos^p\theta\,\sin^q\theta\,d\theta & \text{when } p \text{ is even}\end{cases}$$

The surviving integral is tabulated; we have[8]

$$\int_0^{\frac{1}{2}\pi} \cos^p\theta \,\sin^q\theta\, d\theta = \tfrac{1}{2}B\big(\tfrac{p+1}{2},\tfrac{q+1}{2}\big)$$

where

$$B(x,y) \equiv \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

Here $\Gamma(x) \equiv \int_0^\infty e^{-t}t^{x-1}\,dt$ is Euler's gamma function. Familiarly, $\Gamma(1) = 1$ and $\Gamma(x+1) = x\Gamma(x)$ so when x is an integer one has

$$\Gamma(n+1) = n!$$

from which it follows that

$$\begin{aligned} B(m+1,n+1) &= \frac{\Gamma(m+1)\Gamma(n+1)}{\Gamma(m+n+2)} \\ &= \frac{1}{m+n+1}\cdot\frac{m!n!}{(m+n)!} \end{aligned}$$

Just as Euler's gamma function $\Gamma(x)$ is a function with the wonderful property that at non-negative integral points it reproduces the factorial, so does Euler's beta function $B(x,y)$ possess the property that at non-negative lattice points it reproduces (to within a factor) the combinatorial coefficients. From $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ it follows that at half-integral points one has

$$\Gamma(n+\tfrac{1}{2}) = \frac{(2n-1)!!}{2^n}\sqrt{\pi} = \begin{cases} \frac{1}{2}\sqrt{\pi} & \text{at } n=1 \\ \frac{1\cdot3}{2^2}\sqrt{\pi} & \text{at } n=2 \\ \frac{1\cdot3\cdot5}{2^3}\sqrt{\pi} & \text{at } n=3 \\ \vdots \end{cases}$$

We find ourselves now in position to write

$$\begin{aligned} \int_0^{\pi} \cos^0\theta\,\sin^q\theta\,d\theta &= B\big(\tfrac{1}{2},\tfrac{q+1}{2}\big) \\ &= \frac{\Gamma(\frac{1}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q}{2}+1)} \\ &= \frac{\sqrt{\pi}\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q}{2}+1)} \end{aligned} \tag{29.1}$$

[8] See I. Gradshteyn & I. Ryzhik, *Table of Integrals, Series, and Products* (1965), **3.621.5**, p. 369 or W. Gröbner & N. Hofreiter, *Bestimmte Integrale* (1958), **331.21**, p. 95.

and

$$\begin{aligned}\int_0^\pi \cos^2\theta \sin^q\theta\, d\theta &= B\big(\tfrac{3}{2},\tfrac{q+1}{2}\big)\\ &= \frac{\Gamma(\frac{3}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q}{2}+2)}\\ &= \frac{\frac{1}{2}\sqrt{\pi}\Gamma(\frac{q+1}{2})}{(\frac{q}{2}+1)\Gamma(\frac{q}{2}+1)}\end{aligned} \tag{29.2}$$

It can be shown (and will be shown, though I find it convenient to postpone the demonstration) that

$$S_N(r) = \frac{2\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2})}\, r^{N-1} \tag{30}$$

so

$$\begin{aligned}\frac{S_{N-1}(1)}{S_N(1)} &= \frac{\Gamma(\frac{N}{2})}{\sqrt{\pi}\Gamma(\frac{N-1}{2})}\\ &= \frac{\Gamma(\frac{q}{2}+1)}{\sqrt{\pi}\Gamma(\frac{q+1}{2})} \qquad \text{by } N = q+2\end{aligned} \tag{31}$$

Returning with (29) and (31) to (28) we find, after much cancellation, that $\langle x^0\rangle = 1$ (which is gratifying) and that

$$\langle x^2\rangle = \tfrac{1}{q+2}r^2 = \tfrac{1}{N}r^2$$

Since (26) responds in an obvious way to dimensional generalization, we obtain at once

$$\nabla^2\varphi = \tfrac{2N}{r^2}\big\{\langle\varphi\rangle - \varphi\big\} + \cdots \quad \text{in the } N\text{-dimensional case} \tag{32}$$

which gives back (25) and (27) as special cases. This is a result of (if we can agree to overlook the labor it cost us) some intrinsic charm. But the point to which I would draw my reader's particular attention is this: equation (32) relates a "local" notion—the Laplacian that appears on the left side of the equality—to what might be called a "locally global" notion, for the construction of $\langle\varphi\rangle$ entails *integration* over a (small) *hypersphere*.[9] Both the result and the method of its derivation anticipate things to come.

But before I proceed to my main business I digress again, not just to make myself honest (I have promised to discuss the derivation of (30)) but to plant

[9] "Integration over a hypershpere" is a process which echoes, in a curious way, the "sum over elements of a complete orthonormal set" which at p. 80 in Chapter I entered critically into the *definition* of the Laplacian, as at p. 81 it entered also into the construction of the Poisson bracket.

some more seeds. To describe the volume $V_N(R)$ of an N-sphere of radius R we we might write

$$V_N(R) = \iint\cdots\int_{x_1^2+x_2^2+\cdots+x_N^2=R^2} dx_1 dx_2 \cdots dx_N = V_N \cdot R^N$$

where $V_N = V_N(1)$ is a certain yet-to-be-determined function of N. Evidently the surface area of such a hypersphere can be described

$$S_N(R) = \frac{d}{dR} V_N(R) = S_N \cdot R^{N-1} \quad \text{with} \quad S_N = NA_N$$

Conversely $V_N(R) = \int_0^R S_N(r)\,dr$, which is familiar as the "onion construction" of $V_N(R)$. To achieve the evaluation of S_N—whence of A_N—we resort to a famous trick: pulling

$$I \equiv \iint\cdots\int_{-\infty}^{\infty} e^{-(x_1^2+x_2^2+\cdots+x_N^2)} dx_1 dx_2 \cdots dx_N$$

from our hat, we note that

$$= \begin{cases} \left[\int_{-\infty}^{\infty} e^{-x^2}\,dx\right]^N & \text{on the one hand} \\ S_N \cdot \int_0^{\infty} e^{-r^2} r^{N-1}\,dr & \text{on the other} \end{cases}$$

On the one hand we have the N^{th} power of a Gaussian integral, and from $\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}$ obtain $I = \pi^{N/2}$, while on the other hand we have an integral which by a change of variable (set $r^2 = u$) can be brought to the form $\frac{1}{2}\int_0^{\infty} e^{-u} u^{\frac{N}{2}-1}\,du$ which was seen on p. 85 to define $\Gamma(\frac{N}{2})$. So we have

$$S_N = \frac{2\pi^{\frac{N}{2}}}{\Gamma(\frac{N}{2})}$$

as was asserted at (30), and which entails

$$V_N = \frac{1}{N} S_N = \frac{\pi^{\frac{N}{2}}}{\frac{N}{2}\Gamma(\frac{N}{2})} = \frac{\pi^{\frac{N}{2}}}{\Gamma(\frac{N+2}{2})} \tag{33}$$

Fairly immediately $V_0 = 1$, $V_1 = 2$ and $V_N = \frac{2\pi}{N} V_{N-2}$ so

$$\left.\begin{aligned} V_{N=2n} &= \frac{\pi^n}{n!} \\ V_{N=2n+1} &= 2\pi^n \frac{2^n}{1\cdot 3\cdot 5\cdots(2n+1)} \end{aligned}\right\} \tag{34}$$

which reproduce familiar results at $N = 2$ and $N = 3$. Note the curious fact that one picks up an additional π-factor only at every second step as one advances through ascending N-values.

First steps toward an integral calculus of functionals: Gaussian integration. Look into any text treating "tensor analysis" and you will find elaborate discussion of various derivative structures (covariant derivatives with respect to prescribed affine connections, intrinsic derivatives, constructs which become "accidentally tensorial" because of certain fortuitious cancellations), and of course every successful differentiation process, when "read backwards," becomes a successful antidifferentiation process. But you will find little that has to do with the construction of *definite* integrals. What little you do find invariably depends critically upon an assumption of "total antisymmetry," and falls into the domain of that elegant subject called the "exterior calculus," where integral relations abound, but all are variants of the same relation—called "Stokes' theorem." Or look into any table of integrals. You will find antiderivatives and definite integrals in stupifying (if never quite sufficient) abundance, but very little that pertains to what might be called the "systematics of multiple integration." What little you do find[10] displays a curious preoccupation with hyperspheres, gamma functions, Gaussians. I interpret these circumstances to be symptoms not of neglect but of deep fact: it becomes *possible* to speak in dimensionally extensible generality of the integral properties of multi-variable objects only in a narrowly delimited set of contexts, only in the presence of some highly restrictive assumptions. It would be the business of an "integral calculus of functionals" to assign meaning to expressions of the type

$$\int_{\text{elements of some "function space"}} F[\varphi]\, d[\varphi]$$

and it is sometimes alleged (mainly by persons who find themselves unable to do things they had naively hoped to do) that this is an "underdeveloped subject in a poor state of repair." It is, admittedly, a relatively new subject,[11] but has enjoyed the close attention of legions of mathematicians/physicists of the first rank. My own sense of the situation is that it is not so much lack of technical development as *restrictions inherent in the subject matter* that mainly account for the somewhat claustrophobic feel that tends to attach to the integral calculus of functionals. That said, I turn now to review of a few of the most characteristic ideas in the field.

---

[10] See Gradshteyn & Ryzhik's §§**4.63**–**4.64**, which runs to a total of scarcely five pages (and begins with the volume of an N-sphere!) or the more extended §**3.3** in A. Prudnikov, Yu. Brychkov & O. Marichev's *Integrals and Series* (1986).

[11] N. Wiener was led from the mathematical theory of Brownian motion to the "Wiener integral" only in the 1930s, while R. Feynman's "sum-over-paths formulation of quantum mechanics" dates from the late 1940s.

One cannot expect a multiple integral

$$\int\int\cdots\int F(x^1,x^2,\ldots,x^N)\,dx^1dx^2\cdots dx^N$$

to possess the property of "dimensional extensibility" unless the integrand "separates;" i.e., unless it possesses some variant of the property

$$F(x^1,x^2,\ldots,x^N)=F_1(x^1)F_2(x^2)\cdots F_N(x^N)$$

giving

$$\int\int\cdots\int F(x^1,x^2,\ldots,x^N)\,dx^1dx^2\cdots dx^N=\prod_{i=1}^{N}\int F_i(x^i)\,dx^i$$

If $F(x^1,x^2,\ldots,x^N)=e^{f(x^1,x^2,\ldots,x^N)}$ and $f(x^1,x^2,\ldots,x^N)$ separates in the additive sense $f(x^1,x^2,\ldots,x^N)=f_1(x^1)+f_2(x^2)+\cdots+f_N(x^N)$ then

$$\int\int\cdots\int e^{f(x^1,x^2,\ldots,x^N)}\,dx^1dx^2\cdots dx^N=\prod_{i=1}^{N}\int e^{f_i(x^i)}\,dx^i$$

represents a merely notational variant of the same basic idea, while a more radically distinct variation on the same basic theme would result from

$$F(r^1,\theta^1,r^2,\theta^2,\ldots,r^N,\theta^N)=\prod_i F_i(r^i,\theta^i)$$

"Separation" is, however, a very fragile property in the sense that it is generally not stable with respect to coordinate transformations $\mathbf{x}\longrightarrow\mathbf{y}=\mathbf{y}(\mathbf{x})$; if $F(\mathbf{x})$ separates then $G(\mathbf{y})\equiv F(\mathbf{x}(\mathbf{y}))$ does, in general, *not* separate. Nor is it, in general, easy to discover whether or not $G(\mathbf{y})$ is "separable" in the sense that separation can be achieved by some suitably designed $\mathbf{x}\longleftarrow\mathbf{y}$. A weak kind of stability can, however, be achieved if one writes $F(\mathbf{x})=e^{f(\mathbf{x})}$ and develops $f(\mathbf{x})$ as a multi-variable power series:

$$f(\mathbf{x})=f_0+f_1(\mathbf{x})+\tfrac{1}{2}f_2(\mathbf{x})+\cdots=\sum\tfrac{1}{n!}f_n(\mathbf{x})$$

where $f_n(\mathbf{x})$ is a multinomial of degree n. For if $\mathbf{x}=\mathbf{x}(\mathbf{y})\longleftarrow\mathbf{y}$ is *linear*

$$\mathbf{x}=\mathbb{T}\mathbf{y}+\mathbf{t}$$

then $g_n(\mathbf{y})=f_n(\mathbf{x}(\mathbf{y}))$ contains no terms of degree higher than n, and is itself a multinomial of degree n in the special case $\mathbf{t}=\mathbf{0}$.

Integrals of the type

$$\int\int\cdots\int_{-\infty}^{+\infty}e^{f_0+f_1(\mathbf{x})}\,dx^1dx^2\cdots x^N$$

are divergent, so the conditionally convergent integrals

$$\int\int\cdots\int_{-\infty}^{+\infty} e^{f_0+f_1(\mathbf{x})+\frac{1}{2}f_2(\mathbf{x})}\,dx^1dx^2\cdots x^N$$

acquire in this light a kind of "simplest possible" status. The question is: do they possess also the property of "dimensional extensibility"? Do they admit of discussion in terms that depend so weakly upon N that one can contemplate proceeding to the limit $N\longrightarrow\infty$? They do.

Our problem, after some preparatory notational adjustment, is to evaluate

$$I\equiv\int\int\cdots\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\mathbf{x}\cdot\mathbb{A}\mathbf{x}+2\mathbf{B}\cdot\mathbf{x}+C)}\,dx^1dx^2\cdots x^N$$

To that end, we introduce new variables $\mathbf{y}$ by $\mathbf{x}=\mathbb{R}\mathbf{y}$ and obtain

$$I=\int\int\cdots\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(\mathbf{y}\cdot\mathbb{R}^{\mathsf{T}}\mathbb{A}\mathbb{R}\mathbf{y}+2\mathbf{b}\cdot\mathbf{y}+C)}\,J\,dy^1dy^2\cdots y^N$$

where $\mathbf{b}=\mathbb{R}^{\mathsf{T}}\mathbf{B}$, where the Jacobian $J=\left|\frac{\partial(x^1,x^2,\ldots,x^N)}{\partial(y^1,y^2,\ldots,y^N)}\right|=\det\mathbb{R}$, and where $\mathbb{A}$ can without loss of generality be assumed to be symmetric. Take $\mathbb{R}$ to be in particular the rotation matrix that diagonalizes $\mathbb{A}$:

$$\mathbb{R}^{\mathsf{T}}\mathbb{A}\mathbb{R}=\begin{pmatrix} a_1 & 0 & \cdots & 0\\ 0 & a_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & a_N\end{pmatrix}$$

The numbers a_i are simply the eigenvalues of $\mathbb{A}$; they are necessarily real, and we assume them all to be positive. Noting also that, since $\mathbb{R}$ is a rotation matrix, $J=\det\mathbb{R}=1$, we have

$$I=e^{-\frac{1}{2}C}\cdot\prod_i\underbrace{\int_{-\infty}^{+\infty}e^{-\frac{1}{2}(a_iy^2+2b_iy)}\,dy}$$

$$=\sqrt{\frac{2\pi}{a_i}}\,\exp\left\{\frac{1}{2}\left[b_i\cdot\frac{1}{a_i}\cdot b_i\right]\right\}\tag{35}$$

$$=e^{-\frac{1}{2}C}\cdot\sqrt{\frac{(2\pi)^N}{a_1a_2\cdots a_N}}\,e^{\frac{1}{2}\mathbf{b}\cdot(\mathbb{R}^{\mathsf{T}}\mathbb{A}\mathbb{R})^{-1}\mathbf{b}}$$

$$=e^{-\frac{1}{2}C}\cdot\sqrt{\frac{(2\pi)^N}{\det\mathbb{A}}}\,e^{\frac{1}{2}\mathbf{B}\cdot\mathbb{A}^{-1}\mathbf{B}}\tag{36}$$

Offhand, I can think of no formula in all of pure/applied mathematics that supports the weight of a greater variety of wonderful applications than does the formula that here—fittingly, it seems to me—wears the equation number

$137 = \hbar c/e^2$. One could easily write at book length about those applications, and still not be done. Here I confine myself to a few essential remarks.

Set $\mathbf{B} = -i\mathbf{p}$, $C = 0$ and the multi-dimensional Gaussian Integral Formula (36) asserts that *the Fourier transform of a Gaussian is Gaussian*:

$$\left(\tfrac{1}{\sqrt{2\pi}}\right)^N \int_{-\infty}^{+\infty} e^{i\mathbf{p}\cdot\mathbf{x}}\, e^{-\frac{1}{2}\mathbf{x}\cdot\mathbb{A}\mathbf{x}}\, dx^1 dx^2 \cdots dx^N = \frac{1}{\sqrt{\det \mathbb{A}}} e^{-\frac{1}{2}\mathbf{p}\cdot\mathbb{A}^{-1}\mathbf{p}} \tag{37.1}$$

of which

$$(\det \mathbb{A})^{\frac{1}{4}} e^{-\frac{1}{2}\mathbf{x}\cdot\mathbb{A}\mathbf{x}} \xrightarrow[\text{Fourier transformation}]{} (\det \mathbb{A}^{-1})^{\frac{1}{4}} e^{-\frac{1}{2}\mathbf{p}\cdot\mathbb{A}^{-1}\mathbf{p}} \tag{37.2}$$

provides an even more symmetrical formulation. This result is illustrative of the several "closure" properties of Gaussians. Multiplicative closure—in the sense

$$e^{\text{quadratic}} \cdot e^{\text{quadratic}} = e^{\text{quadratic}}$$

—is an algebraic triviality, but none the less important for that; it lies at the base of the observation that if

$$G(x-m;\sigma) \equiv \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[\frac{x-m}{\sigma}\right]^2 \right\}$$

is taken to notate the familiar "normal distribution" then

$$G(x-m';\sigma') \cdot G(x-m'';\sigma'') = G(m'-m'';\sqrt{\sigma'^2+\sigma''^2}) \cdot G(x-m;\sigma)$$

where $(1/\sigma)^2 = (1/\sigma')^2 + (1/\sigma'')^2$ is less than the lesser of σ' and σ'', while $m = m'(\sigma/\sigma')^2 + m''(\sigma/\sigma'')^2$ has a value intermediate between that of m' and m''. In words:

> normal distribution · normal distribution =
>
> attenuation factor · skinny normal distribution

This result is absolutely fundamental to statistical mechanics, where it gives one license to make replacements of the type $\langle f(x)\rangle \longrightarrow f(\langle x\rangle)$; it is, in short, the reason there exists such a subject as thermodynamics!

Also elementary (though by no means trivial) is what might be called the "integrative closure" property of Gaussians, which I now explain. We proceed from the observation that

$$Q(x) \equiv \begin{pmatrix} x \\ y \end{pmatrix} \cdot \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + 2 \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} + C$$

can be notated

$$Q(x) = ax^2 + 2bx + c$$

with

$$\begin{aligned} a &= A_{11} \\ b &= \tfrac{1}{2}(A_{12} + A_{21})y + B_1 \\ c &= A_{22}y^2 + 2B_2 y + C \end{aligned}$$

Therefore

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}Q(x)}\,dx = \sqrt{\frac{2\pi}{a}}\,\exp\left\{\frac{b^2-ac}{a}\right\}$$

The point is that the result thus achieved has the form

$$= \sqrt{\frac{2\pi}{a}}\,e^{-\frac{1}{2}Q'(y)} \quad \text{with} \quad Q'(y) \equiv a'y^2 + 2b'y + c'$$

so that if one were to undertake a second integration one would confront again an *integral of the original class*.[12] The statement

$$\int_{-\infty}^{+\infty} e^{\text{ quadratic in } N \text{ variables}}\,d(\text{variable}) = e^{\text{ quadratic in } N-1 \text{ variables}} \tag{38}$$

appears to pertain uniquely to quadratics. In evidence I cite the facts that even the most comprehensive handbooks[13] list very few integrals of the type

$$\int_{-\infty}^{+\infty} e^{\text{ non-quadratic polynomial in } x}\,dx$$

and that in the "next simplest case" one has[14]

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^4+2bx^2+c)}\,dx = \sqrt{\tfrac{b}{2a}}e^{\frac{1}{2}\left[\frac{b^2-2ac}{2a}\right]}K_{\frac{1}{4}}\left(\tfrac{b^2}{2a}\right)$$
$$\neq e^{\text{quartic}} \quad !$$

It is "integrative closure" that permits one to construct multiple integrals by tractable iteration of single integrals. The moral appears to be that if it is

---

[12] Indeed, if one works out explicit descriptions of a', b' and c' and inserts them into $\int e^{-\frac{1}{2}Q'(y)\,dy}$ one at length recovers precisely (36), but that would (especially for $N \gg 2$) be "the hard way to go."

[13] See, for example, §§**3.32**–**3.34** in Gradshteyn & Ryzhik and §**2.3.18** in Prudnikov, Brychkov & Marichev.

[14] See Gradshteyn & Ryzhik, **3.323.3**. Properties of the functions $K_\nu(x)$ —sometimes called "Basset functions"—are summarized in Chapter 51 of *An Atlas of Functions* by J. Spanier & K. Oldham (1987). They are constructed from Bessel functions of fractional index and imaginary argument.

iterability that we want, then it is Gaussians that we are stuck with. And Gaussians are, as will emerge, enough.[15]

At the heart of (38) lies the familiar formula

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2+2bx+c)}\,dx = \sqrt{\tfrac{2\pi}{a}}\,e^{\frac{1}{2}\left[\frac{b^2-ac}{a}\right]} \tag{39}$$

of which we have already several times made use. And (39) is itself implicit in this special case:

$$\int_{-\infty}^{+\infty} e^{-x^2}\,dx = \sqrt{\pi} \tag{40}$$

This is a wonderfully simple result to support the full weight of the integral calculus of functionals! We are motivated to digress for a moment to review the derivation of (40), and details of the mechanism by which (40) gives rise to (39). Standardly, one defines $G \equiv \int e^{-x^2}\,dx$ and observes that

$$\begin{aligned} G \cdot G &= \int\int_{-\infty}^{+\infty} e^{-(x^2+y^2)}\,dxdy \\ &= \int_0^{2\pi}\int_0^{\infty} e^{-r^2}\,rdrd\theta \\ &= 2\pi \cdot \tfrac{1}{2}\int_0^{\infty} e^{-s}\,ds \\ &= \pi \end{aligned}$$

from which (40) immediately follows.[16] It is a curious fact that to achieve this result we have had to do a certain amount of "swimming upstream," against the prevailing tide: to evaluate the single integral which lies at the base of our theory of iterative multiple integration we have found it convenient to exploit a change-of-variables trick as it pertains to a certain *double* integral! Concerning

[15] Feynman's "sum-over-paths" is defined by "refinement" $N \longrightarrow \infty$ of just such an iterative scheme. The Gaussians arise there from the circumstance that $\dot{x}$ enters quadratically into the construction of physical Lagrangians. One can readily write out the Lagrangian physics of systems of the hypothetical type $L = \frac{1}{3}\mu\dot{x}^3 - U(x)$. But look at the Hamiltonian: $H = \frac{2}{3}\frac{1}{\sqrt{\mu}}p^{\frac{3}{2}} + U(x)$! Look at the associated Schrödinger equation!! The utter collapse of the Feynman formalism in such cases, the unavailability of functional methods of analysis, inclines us to dismiss such systems as "impossible."

[16] The introduction of polar coordinates entailed tacit adjustment (square to circle) of the domain of integration, which mathematicians correctly point out requires some justification. This little problem can, however, be circumvented

the production of (39) from (40), we by successive changes of variable have

$$\begin{aligned}\sqrt{\tfrac{2\pi}{a}} &= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}ax^2}\,dx \\ &= \int_{-\infty}^{+\infty} e^{-\frac{1}{2}a(x+\frac{b}{a})^2}\,dx \\ &= e^{-\frac{1}{2}\left[\frac{b^2-ac}{a}\right]} \cdot \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(ax^2+2bx+c)}\,dx\end{aligned}$$

What's going on here, up in the exponent, is essentially a "completion of the square." These elementary remarks acquire deeper interest from an observation and a question. Recalling from p. 87 the definition of the gamma function, we observe that

$$\begin{aligned}\int_0^{+\infty} e^{-x^n}\,dx &= \tfrac{1}{n}\int_0^{+\infty} e^{-u}\,u^{\frac{1}{n}-1}\,du \\ &= \tfrac{1}{n}\Gamma(\tfrac{1}{n})\end{aligned}$$

which, by the way, gives back (40) at $n = 2$. Why can we *not* use this striking result to construct a general theory of $\int e^{\text{polynomial of degree } n}\,dx$? Because there exists no cubic, quartic... analog of "completion of the square;"

$$P_n(x) = (x+p)^n + q$$

serves to describe the *most general* monic polynomial of degree n *only in the case* $n = 2$. This little argument provides yet another way (or another face of an old way) to understand that the persistent intrusion of Gaussians into theory of iterative multiple integration (whence into the integral calculus of functionals) is not so much a symptom of "Gaussian chauvinism" as it is a reflection of some essential facts. I have belabored the point, but will henceforth consider it to have been established; we agree to accept, as a fact of life, the thought that if we are to be multiple/functional integrators we are going to have, as a first qualification, to be familiar with all major aspects of *Gaussian* integration theory. Concerning which some important things remain to be said:

---

by a trick which I learned from Joe Buhler: write $y = ux$ and obtain

$$\begin{aligned}G \cdot G &= 2^2 \int_0^\infty \int_0^\infty e^{-x^2(1+u^2)}\,x\,dx\,du \\ &= 4\int_0^\infty \int_0^\infty \frac{e^{-v}}{2(1+u^2)}\,du\,dv \quad \text{where} \quad v = x^2(1+u^2) \\ &= 2\int_0^\infty e^{-v}dv \cdot \int_0^\infty \frac{du}{1+u^2} \\ &= 2 \cdot 1 \cdot \arctan(\infty) \\ &= \pi\end{aligned}$$

Returning to (39), we note that the integrand $e^{-\frac{1}{2}Q(x)}$ will be maximal when $Q(x) \equiv ax^2 + 2bx + c$ is minimal, and from $\frac{d}{dx}Q(x) = 2(ax+b) = 0$ find that $Q(x)$ is minimal at $x = x_0 \equiv -\frac{b}{a}$, where $Q(x_0) = -\left[\frac{b^2-ac}{a}\right]$. The pretty implication is that (39) can be notated

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}Q(x)}\,dx = \sqrt{\tfrac{2\pi}{a}}e^{-\frac{1}{2}Q(x_0)} \tag{41}$$

where

$$Q(x) \equiv ax^2 + 2bx + c \qquad \text{and} \qquad x_0 \equiv -a^{-1}b$$

But by Taylor expansion (which promptly truncates)

$$= Q(x_0) + a(x - x_0)^2$$

Equation (41) can therefore be rewritten in a couple of interesting ways; we have

$$\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\{Q(x_0)+a(x-x_0)^2\}}\,dx = \sqrt{\tfrac{2\pi}{a}}e^{-\frac{1}{2}Q(x_0)} \tag{42.1}$$

—of which more in a minute—and we have the (clearly equivalent) statement

$$\int_{-\infty}^{+\infty} \left\{\sqrt{\tfrac{a}{2\pi}}e^{-\frac{1}{2}a(x-x_0)^2}\right\}\,dx = 1 \quad : \quad \text{all } a \text{ and all } x_0 \tag{42.2}$$

where the integrand is a "normal distribution function," and can in the notation of p. 91 be described $G(x - x_0; a^{-\frac{1}{2}})$. By an identical argument (36) becomes (allow me to write $\mathbf{b}$ and c where formerly I wrote $\mathbf{B}$ and C)

$$\int\int\cdots\int_{-\infty}^{+\infty} e^{-\frac{1}{2}Q(\mathbf{x})}\,dx^1dx^2\cdots x^N = \sqrt{\frac{(2\pi)^N}{\det\mathbb{A}}}\,e^{-\frac{1}{2}Q(\mathbf{x}_0)} \tag{43}$$

where

$$Q(\mathbf{x}) \equiv \mathbf{x}\cdot\mathbb{A}\mathbf{x} + 2\mathbf{b}\cdot\mathbf{x} + c \qquad \text{and} \qquad \mathbf{x}_0 \equiv -\mathbb{A}^{-1}\mathbf{b}$$

By Taylor expansion

$$= Q(\mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)\cdot\mathbb{A}(\mathbf{x} - \mathbf{x}_0)$$

so we have

$$\int\int\cdots\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\{Q(\mathbf{x}_0)+(\mathbf{x}-\mathbf{x}_0)\cdot\mathbb{A}(\mathbf{x}-\mathbf{x}_0)\}}\,dx^1dx^2\cdots x^N = \sqrt{\frac{(2\pi)^N}{\det\mathbb{A}}}\,e^{-\frac{1}{2}Q(\mathbf{x}_0)}$$

and

$$\int\int\cdots\int_{-\infty}^{+\infty} \left\{\sqrt{\frac{\det\mathbb{A}}{(2\pi)^N}}\,e^{-\frac{1}{2}(\mathbf{x}-\mathbf{x}_0)\cdot\mathbb{A}(\mathbf{x}-\mathbf{x}_0)}\right\}\,dx^1dx^2\cdots x^N = 1 \tag{44}$$

The function in braces describes what might be called a "bell-shaped curve in N-space," centered at $\mathbf{x}_0$ and with width characteristics fixed by the eigenvalues of $\mathbb{A}$.

Asymptotic evaluation of integrals by Laplace's method. Pierre Laplace gave the preceeding material important work to do when in 1814 he undertook to study the *asymptotic evaluation of integrals* of the general type

$$I(\lambda) = \int_{-\infty}^{+\infty} f(x)e^{-\lambda g(x)}\,dx$$

Assume $g(x)$ to have a minimum at $x = x_0$. One expects then to have (in ever-better approximation as $\lambda \longrightarrow \infty$)

$$I(\lambda) \sim \int_{x_0-\epsilon}^{x_0+\epsilon} f(x)e^{-\lambda g(x)}\,dx$$

By Taylor expansion $g(x) = g(x_0) + 0 + \frac{1}{2}g''(x_0)(x-x_0)^2 + \cdots$ with $g''(x_0) > 0$. Drawing now upon (42) we obtain

$$I(\lambda) \sim e^{-\lambda g(x_0)}\sqrt{\tfrac{2\pi}{\lambda g''(x_0)}}\int f(x)\left\{\sqrt{\tfrac{\lambda g''(x_0)}{2\pi}}e^{-\frac{1}{2}\lambda g''(x_0)(x-x_0)^2}\right\}dx$$

In a brilliancy which anticipated the official "invention of the δ-function" by more than a century, Laplace observed that the expression in braces nearly vanishes except on a neighborhood of x_0 that becomes ever smaller as λ becomes larger, and arrived thus at the celebrated "Laplace asymptotic evaluation formula"[17]

$$I(\lambda) \sim f(x_0)e^{-\lambda g(x_0)}\sqrt{\tfrac{2\pi}{\lambda g''(x_0)}} \tag{45}$$

In classic illustration of the practical utility of (45) we recall from p. 87 that

$$\Gamma(n+1) \equiv \int_0^\infty e^{-x}x^n\,dx = n! \text{ for } n \text{ integral}$$

But a change of variables $x \longrightarrow y \equiv x/n$ gives

$$= n^{n+1}\int_0^\infty e^{-n(y-\log y)}\,dy$$

and $g(y) \equiv y - \log y$ is minimal at $y = 1$, so by application of (45) we have

$$\Gamma(n+1) = n! \sim \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}$$

[17] For careful discussion of Laplace's formula and its many wonderful variants see Chapter II of A. Erdélyi, *Asymptotic Expansions* (1956) or N. De Bruijn, *Asymptotic Methods in Analysis* (1958).

which is familiar as "Stirling's formula." De Bruijn, in his §4.5, shows how one can, with labor, refine the argument so as to obtain

$$\begin{aligned} n! &\sim \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} \cdot \exp\left\{ \frac{B_2}{1\cdot 2n} + \frac{B_4}{3\cdot 4n^3} + \frac{B_6}{5\cdot 6n^5} + \cdots \right\} \\ &= \sqrt{2\pi} n^{n+\frac{1}{2}} e^{-n} \cdot \left\{ 1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} - \frac{571}{2488320n^4} + \cdots \right\} \end{aligned}$$

where B_2, B_4, B_6, ... are Bernoulli numbers. The pretty particulars of this last result are of less interest than its general implication: Laplace's argument does not simply blurt out its answer and then fall silent; it supports a "refinement strategy" (though this is, to my knowledge, seldom actually used).

I thought I heard some gratuitous coughing during the course of that last paragraph, so hasten to turn now to an "illustration of the practical utility" of Laplace's formula which has a latently more physical feel about it. Let $G(p)$ be the Fourier transform of $F(x)$:

$$F(x) \xrightarrow[\text{Fourier}]{} G(p) = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}px}\, F(x)\, dx$$

Let us, moreover, agree to write $F(x) = \mathcal{F}(x) e^{-\frac{i}{\hbar}f(x)}$ and $G(p) = \mathcal{G}(p) e^{-\frac{i}{\hbar}g(p)}$. The implied relationship

$$\mathcal{G}(p) e^{-\frac{i}{\hbar}g(p)} = \tfrac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathcal{F}(x) e^{-\frac{i}{\hbar}[f(x)-px]}\, dx$$

between $\{\mathcal{F}(x), f(x)\}$ and $\{\mathcal{G}(p), g(p)\}$ is difficult/impossible to describe usefully in general terms, but in the asymptotic limit $\frac{1}{\hbar} \longrightarrow \infty$ we can draw formally upon (45) to obtain

$$\mathcal{G}(p) \cdot e^{-\frac{i}{\hbar}g(p)} \sim \sqrt{\frac{\hbar}{f''(x)}}\, \mathcal{F}(x) \cdot e^{-\frac{i}{\hbar}[f(x)-px]} \Big|_{x\to x(p)} \tag{46}$$

where $x(p)$ is obtained by functional inversion of $p = f'(x)$. The remarkable implication is that $g(p)$ is precisely the *Legendre transform* of $f(x)$! We have established that, in a manner of speaking,

$$\text{Fourier transformations} \quad \longrightarrow \quad e^{\,i\,(\text{Legendre transformations})}$$

and in precisely that same manner of speaking it emerges that

$$\begin{aligned} \text{physical optics} \quad &\xrightarrow[c\longrightarrow\infty]{} \quad e^{\,i\,(\text{geometrical optics})} \\ \text{quantum mechanics} \quad &\xrightarrow[\hbar^{-1}\longrightarrow\infty]{} \quad e^{\,i\,(\text{classical mechanics})} \\ \text{statistical mechanics} \quad &\xrightarrow[k^{-1}\longrightarrow\infty]{} \quad e^{\,-\,(\text{thermodynamics})} \end{aligned}$$

The physical connections thus sketched comprise, I think we can agree, physics of a high order (indeed, physics of an *asymptotically* high order!). Remarkably, at the heart of each of those many-faceted connections live either Laplace's formula or one of its close relatives (the Riemann–Debye "method of steepest descents," the Stokes–Kelvin "method of stationary phase"). And at the heart of each of those lives a Gaussian integral.[18]

Laplace's asymptotic formula admits straightforwardly of N-dimensional generalization. We write

$$I(\lambda) \equiv \int\int\cdots\int_{-\infty}^{+\infty} F(\mathbf{x})e^{-\lambda g(\mathbf{x})}\ dx^1dx^2\cdots dx^N$$

Assume $g(\mathbf{x})$ to have a minimum at $\mathbf{x} = \mathbf{x}_0$. Then

$$g(\mathbf{x}) = g(\mathbf{x}_0) + 0 + \tfrac{1}{2}(\mathbf{x}-\mathbf{x}_0)\cdot\mathbb{G}(\mathbf{x}-\mathbf{x}_0) + \cdots$$

where $\mathbb{G} \equiv \|\partial^2 g(\mathbf{x})/\partial x^i \partial x^j\|$ —the matrix of second partials, evaluated at $\mathbf{x}_0$ —is positive definite. Arguing as before, we obtain

$$I(\lambda) \sim F(\mathbf{x}_0)e^{-\lambda g(\mathbf{x}_0)}\sqrt{\frac{(2\pi/\lambda)^N}{\det\mathbb{G}}} \tag{47}$$

Physically motivated functional integration. I turn now to discussion of how a theory of functional integration emerges " by refinement" (i.e., in the limit $N \to \infty$) from the interative theory of multiple integration. Both Wiener (in the late 1920's, for reasons characteristic of his approach to the theory of Brownian motion[19]) and Feynman (in the early 1940's, for reasons characteristic of his approach to quantum mechanics[20]) had reason to be interested in what have come universally (if awkwardly) to be called "sum-over-path" processes.

[18] I remarked in the text that (46) was obtained by "formal" application of (45). The adjective alludes to the fact that the Gaussian integral formula (39) holds if an only if $\Re(a) > 0$, which in the present context may not be satisfied. The problem would not have arisen had we been discussing Laplace transforms rather than Fourier transforms, and can frequently be circumvented by one or another of strategies which physicists have been at pains to devise; for example, one might (as Feynman himself suggested: see footnote #13 to his "Space-time approach to non-relativistic quantum mechanics," Rev. Mod. Phys. **20**, 367 (1948)) make the replacement $\hbar \longrightarrow \hbar(1 - i\epsilon)$ and then set $\epsilon \downarrow 0$ at the end of the day. My own practice will be to proceed with formal abandon, trusting to the sensible pattern of our (formal) results, and to the presumption that when we have accumulated results in a sufficient mass we will find both motivation and some elegant means to dot the i's and cross the mathematical t's.

[19] See Chapter I of his *Nonlinear Problems in Random Theory* (1958).

[20] See §4 of the classic paper cited on the previous page, or Chapter II of *Quanatum Mechanics and Path Integrals* by R. Feynman & A. Hibbs (1965).

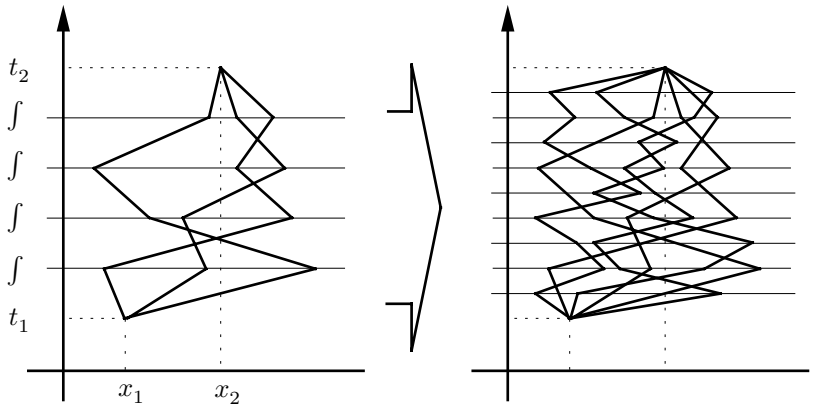


FIGURE 2: *Representation of the elementary essentials of the idea from which the Wiener–Feynman "sum-over-paths" construction proceeds.*

Each worked in a subject area marked (as it happens) by the natural occurrence —for distinct reasons—of Gaussians, each was led to contemplate expressions of the type

$$\lim_{N\to\infty} \int \cdots \int e^{-(\text{quadratic form in N variables})}\, d(\text{variables}) \tag{48}$$

and each was protected from disaster by the "integrative closure" property of Gaussians. Each was led to write something like

$$\int_{\text{space of paths } x(t)} F[x(t)]\mathcal{D}x(t) \tag{49}$$

to describe the result of such a limiting process. Concerning the structure of the "space of paths" over which the functional integral (49) ranges: the figures suggests that the elements $x(t)$ of "path space" are, with rare exceptions, too spiky to permit the construction of $\dot{x}(t)$. It would, however, be a mistake to waste time pondering whether this development is to be regarded as a "physical discovery" or a "formal blemish," for to do so would be to attach to the figure a literalness it is not intended to support. Suppose, for example, we were to write

$$x(t) = x_{\text{nice}}(t) + s(t)$$

where $x_{\text{nice}}(t)$ is *any* (nice or unnice) designated path linking specified spacetime endpoints $(x_1, t_1) \longrightarrow (x_2, t_2)$ and where

$$s(t) \equiv \sum_{n=1}^{\infty} a_n \sin\left\{ n\pi \left[\frac{t\ - t_1}{t_2 - t_1} \right] \right\} \tag{50}$$

has by design the property that $s(t_1) = s(t_2) = 0$. Individual paths would then be specified not by "the locations of their kinks" but by their Fourier coefficients $\{a_n\}$. Elements of the path space thus constructed can be expected to have differentiability properties quite different from those originally contemplated, and "summing-over-paths" would entail iterated operations of the type $\int da$.[21] Applications of the functional integral concept tend, to a remarkable degree, to proceed independently of any precise characterization of path space.

In order to make as concretely clear as possible the issues and methods most characteristic of the applied integral calculus of functionals, I look now to the essential rudiments of the Feynman formalism. By way of preparation, in order to grasp Feyman's train of thought, we remind remind ourselves that in abstract quantum mechanics one has $i\hbar\frac{\partial}{\partial t}|\psi) = \mathsf{H}\,|\psi)$, giving $|\psi)_t = \exp\{\frac{1}{i\hbar}\mathsf{H}t\}|\psi)_0$. In the x-representation[22] we have $(x|\psi)_t = \int (x|\exp\{\frac{1}{i\hbar}\mathsf{H}t\}|y)dy(y|\psi)_0$ which is more often written

$$\psi(x,t) = \int K(x,t;y,0)\psi(y,0)\,dy$$

It is from the preceeding equaton that the Green's function of the Schrödinger equation—usually called the "propagator"

$$K(x,t;y,0) \equiv (x,t|y,0) = (x|\exp\{\tfrac{1}{i\hbar}\mathsf{H}t\}|y)$$

—acquires its role as the "fundamental object of quantum dynamics." Three properties of the propagator are of immediate importance. We note first that $K(x,t;\bullet,\bullet)$ is itself a solution of the Schrödinger equation

$$i\hbar\tfrac{\partial}{\partial t}K(x,t;\bullet,\bullet) = \mathsf{H}\,K(x,t;\bullet,\bullet) \tag{51.1}$$

From

$$\lim_{t\downarrow 0} K(x,t;y,0) = (x|y) = \delta(x-y) \tag{51.2}$$

we see that $K(x,t;y,0)$ is in fact the solution that evolved from an initial δ-function. It follows finally from the triviality $e^{\mathsf{H}(a+b)} = e^{\mathsf{H}a}\cdot e^{\mathsf{H}b}$ that

$$K(x,t;z,0) = \int K(x,t;y,\tau)dyK(y,\tau;y,0) \quad \text{for all } t \geq \tau \geq 0 \tag{51.3}$$

It was by an interative refinement procedure based upon the "composition rule" that Feynman was led to the imagery of FIGURE 10. But it was a stroke of

[21] For discussion of details relating to this mode of proceeding, see Chapter I, pp. 56–60 of QUANTUM MECHANICS (1967).

[22] The "space-time approach..." of Feynman's title reflects his appreciation of the fact that selection of the x-representation is an arbitrary act, yet an act basic to the imagery from which his paper proceeds.

genius[23] which led Feynman to contemplate a formula of the structure

$$K(x_2,t_2;x_1,t_1) = \int e^{\frac{i}{\hbar}S[x(t)]}\mathcal{D}x(t) \tag{52}$$

Here $x(t)$ is a "path" with the endpoint properties

$$x(t_1) = x_1 \quad \text{and} \quad x(t_2) = x_2$$

$S[x(t)]$ is the *classical* action functional associated with that path

$$S[x(t)] = \int_{t_1}^{t_2} L(x(t),\dot{x}(t))\,dt \tag{53}$$

and $\mathcal{D}x(t)$ —for which in some contexts it becomes more natural to write $R[x(t)]\mathcal{D}x(t)$ —alludes implicitly to the as-yet-unspecified "measure-theoretic" properties of path space. Our problem is to assign specific meaning to the functional integral that stands on the right side of (52). To that end, let $L(x,\dot{x}) = \frac{1}{2}m\dot{x}^2 - U(x)$ describe the classical dynamics of some one-dimensional system, let $x_c(t)$ be a solution of the equations of motion that interpolates $(x_1,t_1) \longrightarrow (x_2,t_2)$ between specified endpoints, let $s(t)$ be some given/fixed nice function with the property that $s(t_1) = s(t_2) = 0$ and let

$$x(t) = x_c(t) + \lambda s(t)$$

be the elements of a *one-parameter path space generated by* $s(t)$. Under such circumstances the action functional (53)—though it remains a functional of $s(t)$ —becomes an ordinary function of the parameter λ (and of the endpoint coordinates). This is the simplification that makes the present discussion[24] work. We have

$$\begin{aligned} L(x_c+\lambda s,\dot{x}_c+\lambda\dot{s}) &= e^{\lambda\left(s\frac{\partial}{\partial x_c}+\dot{s}\frac{\partial}{\partial \dot{x}_c}\right)}L(x_c,\dot{x}_c) \\ &= \sum_k \tfrac{1}{k!}\lambda^k L_k(x_c,\dot{x}_c,s,\dot{s}) \end{aligned}$$

giving

$$\begin{aligned} S[x(t)] = \sum_k \tfrac{1}{k!}\lambda^k \underbrace{S_k(x_2,t_2;x_1,t_1;s(t))}& \\ &= \int_{t_1}^{t_2} L_k(x_c.\dot{x}_c,s,\dot{s})\,dt \end{aligned}$$

---

[23] *Dirac's* genius, one might argue. See §32 "The action principle," in *The Principles of Quantum Mechanics* (1958) and "The Lagrangian in quantum mechanics," Physik. Zeits. Sowjetunion **3**, 64 (1933), both of which—and little else—are cited by Feynman. The latter paper has been reprinted in J. Schwinger (ed.) *Quantum Electrodynamics* (1958).

[24] It has served my expository purpose to depart here from the historic main line of Feynman's argument; I follow instead in the footsteps of C. W. Kilmister, "A note on summation over Feynman histories," Proc. Camb. Phil. Soc. **54**, 302 (1958).

and notice that

$$\begin{aligned} S_0 &= S[x_c(t)] \quad \text{is just the CLASSICAL ACTION} \\ S_1 &= 0 \qquad\qquad \text{by HAMILTON'S PRINCIPLE} \end{aligned}$$

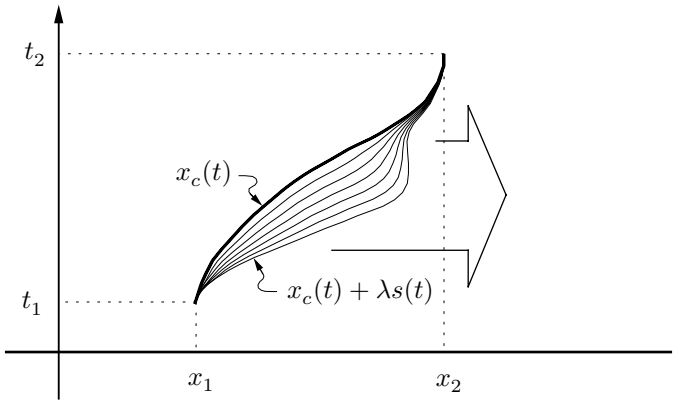


FIGURE 3: *λ-parameterized family of paths having $x_c(t)$ as a member, and generated by an arbitrary $s(t)$. The arrow indicates the effect of increasing λ. We are making integral use of what is, in fact, the construction standard to the calculus of variations.*

By computation

$$\begin{aligned} L_0 &= \tfrac{1}{2} m \dot{x}_c^2 - U(x_c) \\ L_1 &= \text{need not be computed} \\ L_2 &= \quad m\,\dot{s}^2 - U''(x_c)\,s^2 \\ &\;\;\vdots \\ L_k &= \qquad\quad - U^{(k)}(x_c)\,s^k \end{aligned}$$

so within the path space here in question we have[25]

$$\begin{aligned} S[x(t)] = S_{\text{classical}}(x_2, t_2; x_1, t_1) &+ \tfrac{1}{2}\lambda^2 \int_{t_1}^{t_2} \left\{ m\,\dot{s}^2 - U''(x_c) s^2 \right\} dt \\ &- \sum_{k=3}^{\infty} \tfrac{1}{k!} \lambda^k \int_{t_1}^{t_2} U^{(k)}(x_c)\, s^k \, dt \end{aligned} \tag{54}$$

[25] This is a specialized instance of (see again (9)) the generic Volterra series

$$\begin{aligned} S[x_c + \lambda s] = S[x_c] &+ \lambda \int_{t_1}^{t_2} \frac{\delta S[x_c]}{\delta x_c(t)}\, s(t)\, dt \\ &+ \tfrac{1}{2}\lambda^2 \int_{t_1}^{t_2}\int_{t_1}^{t_2} \frac{\delta^2 S[x_c]}{\delta x_c(t')\delta x_c(t'')}\, s(t') s(t'')\, dt' dt'' + \cdots \end{aligned}$$

It becomes natural at this point to write

$$\int_{\text{paths generated by } s(t)} e^{\frac{i}{\hbar}S[x(t)]}\,\mathcal{D}x(t) = \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}\{S_0+\frac{1}{2}\lambda^2 S_2+\text{higher order terms}\}}\,d\lambda \qquad (55)$$

In the interests of maximal tractability (always fair in exploratory work) we opt to kill the "higher order terms" by assuming the potential $U(x)$ to depend at most quadratically upon x; we assume, in short, that the Lagrangian $L(x,\dot{x})$ pertains to an *oscillator in a gravitational field*:

$$U(x) = mgx + \tfrac{1}{2}m\omega^2x^2 \qquad (56)$$

Equation (54) then truncates:

$$S[x(t)] = \underbrace{S_{\text{classical}}(x_2,t_2;x_1,t_1)}_{S_0} + \tfrac{1}{2}\lambda^2\cdot\underbrace{\int_{t_1}^{t_2} m\{\dot{s}^2-\omega^2s^2\}\,dt}_{S_2} \qquad (57)$$

Remarkably, all reference to $x_c(t)$—and therefore to the variables x_1 and x_2—has disappeared from the 2nd-order term, about which powerful things of several sorts can be said. We might write

$$S_2 = S_2[s(t)] = \mathbf{D}^2_{[s]}S[x(t)] \qquad (58.1)$$

to emphasize that S_2 is a functional of $s(t)$ from which all $x(t)$-dependence has dropped away. And we might write

$$S_2 = S_2(t_2-t_1) \qquad (59.2)$$

to emphasize that S_2 depends upon t_1 and t_2 only through their difference,[26] and is (as previously remarked) independent of x_1 and x_2 . If we now return

[26] It is elementary that

$$\int_{t_1}^{t_2} F\big(x(\tfrac{t\ -t_1}{t_2-t_1}),\dot{x}(\tfrac{t\ -t_1}{t_2-t_1})\big)\,dt = \int_0^1 F\big(x(\vartheta),\tfrac{1}{t_2-t_1}\tfrac{d}{d\vartheta}x(\vartheta)\big)\big(t_2-t_1\big)d\vartheta$$
$$= \text{function of } (t_2-t_1)$$

so one has only to insert (50) into the integral that defines S_2 to achieve the result claimed in the text. One could, however, continue; drawing upon

$$\int_0^1 \sin m\pi\vartheta\sin n\pi\vartheta\,d\vartheta = \int_0^1 \cos m\pi\vartheta\cos n\pi\vartheta\,d\vartheta = \tfrac{1}{2}\delta_{mn}$$

for $m,n = 1,2,3,\cdots$ one can actually *do* the integral. One obtains at length

$$S_2 = \tfrac{m}{2T}\sum(\pi n)^2\Big[1-\big(\tfrac{\omega T}{\pi n}\big)^2\Big]a_n^2 \quad \text{with} \quad T\equiv t_2-t_1$$

which is the starting point for the $\int da$-procedure to which I alluded on p. 99.

with (57) to (55) we obtain

$$\int_{\text{paths generated by } s(t)} e^{\frac{i}{\hbar}S[x(t)]}\,\mathcal{D}x(t) = e^{\frac{i}{\hbar}S_0}\cdot\int_{-\infty}^{+\infty} e^{-\frac{1}{2}\cdot\frac{1}{i\hbar}S_2[s(t)]\cdot\lambda^2} \tag{59}$$

To perform the Gaussian integral is to obtain

$$= e^{\frac{i}{\hbar}S_0}\cdot\sqrt{\frac{2\pi i\hbar}{S_2[s(t)]}}$$

which, because of its surviving functional dependence upon the arbitrarily selected generator $s(t)$, cannot possibly provide a description of the propagator $(x_2,t_2|x_1,t_1)$. One obvious way to remedy this defect is—consistently with the essential spirit of the Feyman formalism—to *sum over all generators*; we back up to (59), set $\lambda=1$, and obtain

$$\begin{aligned} K(x_2,t_2;x_1,t_1) &= \int_{\text{all paths}} e^{\frac{i}{\hbar}S[x(t)]}\mathcal{D}x(t) \\ &= e^{\frac{i}{\hbar}S_0(x_2,t_2;x_1,t_1)}\cdot\left\{\int_{\text{all generators}} e^{\frac{i}{\hbar}\frac{1}{2}S_2[s(t)]}\mathcal{D}s(t)\right\} \end{aligned}$$

We appear to have simply replaced one functional integral by another, but the latter is an object we know something about: it is (since a sum of such functions) a function of t_2-t_1. So we have

$$K(x_2,t_2;x_1,t_1) = A(t_2-t_1)\cdot e^{\frac{i}{\hbar}S_0(x_2,t_2;x_1,t_1)} \tag{60.1}$$

with

$$A(t_2-t_1) = \int_{\text{all generators}} e^{\frac{i}{\hbar}\frac{m}{2}\int_{t_1}^{t_2}\{\dot{s}^2-\omega^2 s^2\}\,dt}\mathcal{D}s(t) \tag{60.2}$$

There are several alternative ways in which we might now proceed. We might roll up our sleeves and undertake (as Feynman did) to *evaluate* the functional integral that defines $A(t_2-t_1)$. To that end we would write

$$\begin{aligned} &\int_{t_1}^{t_2}\{\dot{s}^2-\omega^2 s^2\}\,dt \\ &\quad= \lim_{N\to\infty}\tau\Big\{\big(\tfrac{s_1-0}{\tau}\big)^2+\big(\tfrac{s_2-s_1}{\tau}\big)^2+\cdots+\big(\tfrac{s_N-s_{N-1}}{\tau}\big)^2+\big(\tfrac{0-s_N}{\tau}\big)^2 \\ &\qquad\qquad -\omega^2\big(s_1^2+s_2^2+\cdots+s_{N-1}^2+s_N^2\big)\Big\} \\ &\quad= \lim_{N\to\infty}\tfrac{1}{\tau}\mathbf{s}\cdot\mathbb{M}\mathbf{s} \end{aligned}$$

with $\tau=(t_2-t_1)/(N+1)=[(t_2-t_1)/N]\{1-\frac{1}{N}+\frac{1}{N^2}+\cdots\}\sim(t_2-t_1)/N$ and

$$\mathbb{M} \equiv \begin{pmatrix} M & -1 & 0 & 0 & & & 0 \\ -1 & M & -1 & 0 & & & 0 \\ 0 & -1 & M & -1 & & & 0 \\ & & & \ddots & & & \\ & & & & \ddots & & \\ 0 & & & -1 & M & -1 & \\ 0 & & & 0 & -1 & M & \end{pmatrix} \quad \text{where} \quad M \equiv 2 - (\tau\omega)^2$$

We look to the N-fold integral

$$I_N = \int \cdots \int_{\infty}^{+\infty} e^{-\frac{1}{2}\mathbf{s}\cdot\mathbb{A}\mathbf{s}}\, ds_1 ds_2 \cdots ds_N$$

where $\mathbb{A} = \beta\mathbb{M}$ and $\beta = m/i\hbar\tau$ and draw upon the multi-dimensional Gaussian integral formula to obtain

$$I_N = \sqrt{\frac{(2\pi)^N}{\det \mathbb{A}}} = \sqrt{\frac{(2\pi/\beta)^N}{\det \mathbb{M}}} \tag{61}$$

To evaluate $D_N = \det \mathbb{M}$ when $\mathbb{M}$ is $N \times N$, we look to the sequence of

$$D_1 = (\, M \,), \quad D_2 = \begin{pmatrix} M & -1 \\ -1 & M \end{pmatrix}, \quad D_3 = \begin{pmatrix} M & -1 & 0 \\ -1 & M & -1 \\ 0 & -1 & M \end{pmatrix}, \quad \ldots$$

of sub-determinants and obtain

$$\begin{aligned} D_1 &= M \\ D_2 &= M^2 - 1 \\ D_3 &= M^3 - 2M \\ &\vdots \\ D_n &= MD_{n-1} - D_{n-2} \end{aligned}$$

We introduce the "tempered" numbers $\mathcal{D}_n \equiv \omega\tau D_n$ to facilitate passage to the limit. They obviously satisfy an identical recursion relation, and upon recalling the definition of M we observe that the recursion relation in question can be expressed

$$\frac{1}{\tau}\left\{ \frac{\mathcal{D}_n(N) - \mathcal{D}_{n-1}(N)}{\tau} - \frac{\mathcal{D}_{n-1}(N) - \mathcal{D}_{n-2}(N)}{\tau} \right\} = -\omega^2 \mathcal{D}_{n-1}(N)$$

This in the limit becomes a differential equation

$$\frac{d^2\mathcal{D}(t)}{dt^2} = -\omega^2 \mathcal{D}(t) \tag{62.1}$$

descriptive of a function $D(t)$ for which we seek the value at $t = N\tau = t_2 - t_1$. To start the recursive construction off we need initial data; we have

$$\mathcal{D}_1 = \omega\tau[2 - (\omega\tau)^2] \quad \text{giving} \quad \mathcal{D}(0) = 0 \tag{62.2}$$

and

$$\begin{aligned}\frac{\mathcal{D}_2 - \mathcal{D}_1}{\tau} &= \frac{\omega\tau}{\tau}\Big\{[2 - (\omega\tau)^2]^2 - 1 - [2 - (\omega\tau)^2]\Big\} \\ &= \omega\Big\{1 - 3(\omega\tau)^2 + (\omega\tau)^4\Big\} \quad \text{giving} \quad \mathcal{D}'(0) = \omega\end{aligned} \tag{62.3}$$

It follows from (62) that $\mathcal{D}(t) = \sin\omega t$. Returning with this information to (61) we obtain

$$\begin{aligned}I_N &= \left(\frac{2\pi i\hbar\tau}{m}\right)^{\frac{N}{2}} \sqrt{\frac{\omega\tau}{\sin\omega(t_2 - t_1)}} \\ &= R^{N+1} \cdot \sqrt{\frac{m\omega}{2\pi i\hbar \sin\omega(t_2 - t_1)}} \quad \text{with} \quad R(\tau) \equiv \sqrt{\frac{2\pi i\hbar\tau}{m}}\end{aligned}$$

To obtain a non-trivial result in the limit $\tau \downarrow 0$ we must abandon the prefactor. To that end we make the replacement

$$ds_1 d_2 \cdots d_N \longrightarrow R \cdot ds_1 \cdot R \cdot ds_2 \cdot R \cdots R \cdot ds_N \cdot R$$

which is, in effect, to assign a "measure" to path space. Thus—following a cleaned-up version of the path blazed by Feynman—do we assign direct meaning to the statement

$$A(t_2 - t_1) = \int_{\text{all generators}} e^{\frac{i}{\hbar}\frac{m}{2}\int_{t_1}^{t_2}\{\dot{s}^2 - \omega^2 s^2\}\,dt} \mathcal{D}s(t) = \sqrt{\frac{m\omega}{2\pi i\hbar \sin\omega(t_2 - t_1)}} \tag{63}$$

Our success, it will be noted, was entirely Gaussian in origin. And hard won!

There is, however, a "softer" way to proceed. We might consider that the functional integral concept had already told us what it had to say when at (60.1) it ascribed a certain non-obvious *structure* to the propagator, and that it is to conditions (51) that we should look for more particular information about the left side $A(t_2 - t_1)$. To illustrate with minimal clutter the kind of analysis I have in mind, consider the case of a free particle. For such a system it is a familiar fact that the classical action can be described

$$S_0(x_2, t_2; x_1, t_1) = \frac{m}{2}\frac{(x_2 - x_1)^2}{t_2 - t_1}$$

What condition on $A(t_2 - t_1)$ is forced by the requirement that, consistently with (51.2),

$$K(x_2, t_2; x_1, t_1) = A(t_2 - t_1) \cdot e^{\frac{i}{\hbar}\frac{m}{2}(x_2 - x_1)^2/(t_2 - t_1)} \quad \longrightarrow \quad \delta(x_2 - x_1)$$

as $(t_2 - t_1) \downarrow 0$? Familiarly

$$\delta(x-a) = \lim_{\sigma \to 0} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[\frac{x-a}{\sigma}\right]^2 \right\}$$

so we write

$$e^{\frac{i}{\hbar}\frac{m}{2}(x_2-x_1)^2/(t_2-t_1)} = \exp\left\{ -\frac{1}{2}\left[\frac{x_2-x_1}{\sqrt{i\hbar(t_2-t_1)/m}}\right]^2 \right\}$$

and conclude that $A(t_2 - t_1)$ has necessarily the form

$$A(t_2-t_1) = \sqrt{\frac{m}{2\pi\hbar(t_2-t_1)}} \cdot \{1 + \text{arbitrary power series in } (t_2-t_1)\}$$

This result is consistent with the result obtained from (63) in the free particle limit $\omega \downarrow 0$. Much sharper conclusions can be drawn from (51.3); one wants

$$\begin{aligned} &A(t_2-t_1) \cdot e^{\frac{i}{\hbar}\frac{m}{2}(x_2-x_1)^2/(t_2-t_1)} \\ &\quad = A(t_2-t)A(t-t_1) \int_{-\infty}^{+\infty} e^{\frac{i}{\hbar}\frac{m}{2}(x_2-x)^2/(t_2-t)} e^{\frac{i}{\hbar}\frac{m}{2}(x-x_1)^2/(t-t_1)}\, dx \end{aligned}$$

which after performance of the Gaussian integral is found to entail

$$A(t_2-t_1) = A(t_2-t)A(t-t_1)\sqrt{\frac{2\pi i\hbar}{m}\frac{(t_2-t)(t-t_1)}{(t_2-t_1)}}$$

It implication is that $A(\bullet)$ satisfies a functional equation of the form

$$A(x+y) = A(x)A(y)\sqrt{\frac{\alpha x \cdot \alpha y}{\alpha(x+y)}} \quad \text{with} \quad \alpha \equiv \sqrt{\frac{2\pi i\hbar}{m}}$$

This can be written $G(x+y) = G(x)G(y)$ with $G(x) \equiv A(x)\sqrt{\alpha x}$, and if $\Gamma(x) \equiv \log G(x)$ we have $\Gamma(x+y) = \Gamma(x) + \Gamma(y)$. Therefore

$$\frac{\Gamma(x+y) - \Gamma(x)}{y} = \frac{\Gamma(y)}{y} \quad \text{for all } x$$

from which (taking Y to the limit $y \downarrow 0$) we obtain

$$\frac{d\Gamma(x)}{dx} = k \quad \text{whence} \quad \Gamma(x) = kx + c$$

But the functional condition satisfied by $\Gamma(x)$ enforces $c = 0$, so we have $G(x) = e^{kx}$ giving $A(x) = \sqrt{\frac{1}{\alpha x}}e^{kx}$. Thus do we obtain

$$K_{\text{free particle}} = \sqrt{\frac{m}{2\pi i\hbar(t_2-t_1)}} \cdot e^{k(t_2-t_1)} \exp\left\{ \tfrac{i}{\hbar} S_{\text{free particle}} \right\}$$

The central exponential can be gauged away by adjustment of the energy scale, since $V(x) \longrightarrow V(x) + k$ induces $S \longrightarrow S - k(t_2 - t_1)$. This result is not only consistent with, but actually reproduces, the result implicit in (63). What information could, alternatively, have been gleaned from the requirement (51.1)that $K(x,t;\bullet,\bullet) = A(t)\exp\{\frac{i}{\hbar}S(x,t;\bullet,\bullet)\}$ satisfy the Schrödinger equation? The first of the equations (68) reduces (by $A_{xx} = 0$) to precisely the Hamilton-Jacobi equation, which S by construction satisfies exactly. The second of equations (68) in Chapter I reads

$$(A^2)_t + (\tfrac{x-x_1}{t-t_1}A^2)_x = 0$$

giving

$$A(t-t_1) = A(t_0)\cdot\sqrt{\frac{t_0}{t\ \ -t_1}}$$

which is again consistent with but much weaker than the result implicit in (63). The evidence of this discussion suggests that after-the-fact constructions of $A(t_2 - t_1)$ proceed most effectively from the composition rule (51.3).

And there is, in fact, a still "softer" line of argument which is sufficient to the needs of some applications. Looking back again to (60.1), we note that *ratios* of propagators are described by a formula

$$\frac{K(x_2,t_2;x_1,t_1)}{K(\tilde{x}_2,t_2;\tilde{x}_1,t_1)} = e^{\frac{i}{\hbar}\{S(x_2,t_2;x_1,t_1)-S(\tilde{x}_2,t_2;\tilde{x}_1,t_1)\}}$$

from which all reference to $A(t_2 - t_1)$ has dropped away. This result becomes most vivid when $\tilde{x}_1$ is a "vacuum point"—a point at which the particle can be *at rest with zero energy*; we have

$$S(\tilde{x},t_2;\tilde{x},t_1) = \begin{cases} -E_0(t_2-t_1) & \text{when } \tilde{x} \text{ is an equilibrium point} \\ 0 & \text{when } \tilde{x} \text{ is a "vacuum point"} \end{cases}$$

and in the latter case

$$K(x_2,t_2;x_1,t_1) = K(\tilde{x},t_2;\tilde{x},t_1)\cdot e^{\frac{i}{\hbar}S(x_2,t_2;x_1,t_1)} \tag{64}$$

For a free particle

$$S_{\text{free particle}}(x_2,t_2;x_1,t_1) = \frac{m}{2}\frac{(x_2-x_1)^2}{t_2-t_1}$$

shows that *every* point is a "vacuum point":

$$S_{\text{free particle}}(x,t_2;x,t_1) = 0 \quad \text{for all } x$$

For an oscillator

$$S_{\text{oscillator}}(x,t_2;x,t_1) = \frac{m\omega}{2\sin\omega(t_2-t_1)}\Big[(x_2^2+x_1^2)cos\omega(t_2-t_1) - 2x_2x_1\Big]$$

there is a single vacuum point, situated at the origin. For the system

$$V(x) = mgx + \tfrac{1}{2}m\omega^2 x^2$$

a solitary equilibrium point resides at $\tilde{x} = -g/\omega^2$, where the rest energy is $E_0 = -mg^2/\omega^2$; to make $\tilde{x}$ into a vacuum point one must adjust the zero of the energy scale. For a particle in free fall $V(x) = mgx$ there is *no* equilibrium point, no vacuum point, and it becomes therefore impossible to make use of (64).

The Feynman formalism—clearly and explicitly—takes classical mechanics as its point of departure, and achieves quantum mechanics by a functional integration process, a process that (see the figure) "gives Hamilton's comparison

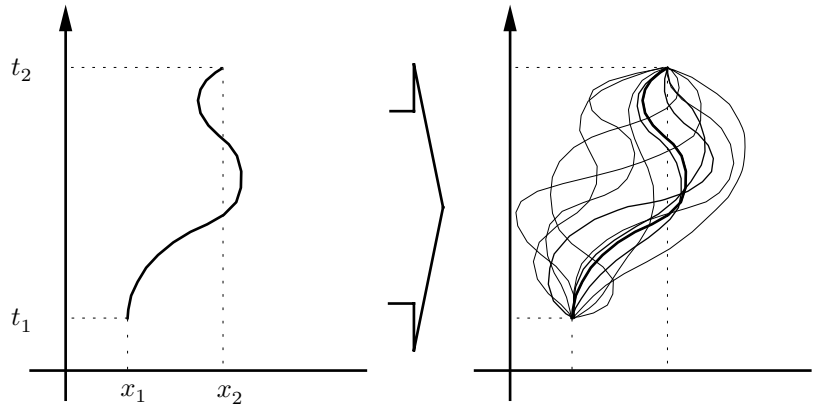


FIGURE 4: *Representation of the Feynman quantization procedure.*

paths a physical job to do." It is, in effect, a *quantization procedure*, and is today widely considered to be "the quantization procedure of choice." Run in reverse, it provides fresh insight into the placement of classical mechanics within a quantum world, and it is that aspect of the Feynman formalism that I want now to explore. Let

DIRAC EQUATION

in 2-dimensional spacetime

Nicholas Wheeler, Reed College Physics Department
January 2000

Introduction. Expository special relativity is, for many purposes, well served by the pretense that we inhabit a spacetime of $1+1$ dimensions, though in such a world proper rotation is impossible, boosts are necessarily colinear, and certain kinematic phenomena—Thomas precession, most notably—remain hidden from view.

The Schrödinger equation came into the world wearing 3-dimensional dress, intent upon serious physical business (physics of the hydrogen atom); the "quantum theory of 1-dimensional systems" was a pedagogical afterthought. Similarly serious were Dirac's motivating objectives (theory of the electron, relativistically corrected theory of hydrogen, clarification of spin concept), but in the latter instance the 3-dimensionality of space has seemed so central to the architecture of the theory that no tradition of "pedagogical pull-back to lower dimension" has come into being. I am motivated to inquire into the question of whether or not such a "toy Dirac theory" has things to teach us. We will find that it has, in fact, *many* valuable lessons to impart, and that it speaks of deep things with engaging simplicity.

Suppose, therefore, that we are relativistically informed one-dimensional physicists who, in reference to our inertial frame (and in the absence of gravity), write

$$\boldsymbol{g} \equiv \|g_{\mu\nu}\| \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{1}$$

to describe the Lorentzian metric structure of spacetime. It has been our habit to erect our theory of the wave equation $\square\psi = 0$ on a "factorization trick"

$$\square \equiv g^{\mu\nu}\partial_\mu\partial_\nu = \partial_0^2 - \partial_1^2 = (\partial_0 - \partial_1)(\partial_0 + \partial_1)$$

and we have noticed, by the way, that the trick fails when we attempt to apply it to the Klein-Gordon equation $(\square + \varkappa^2)\psi = 0$.

Proceeding in imitation of Dirac, we observe that the wave operator can be written as a square

$$\square = (\boldsymbol{\gamma}^\mu\partial_\mu)(\boldsymbol{\gamma}^\nu\partial_\nu) \tag{2}$$

provided $\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu + \boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\mu = 2g^{\mu\nu}\mathbf{I}$, of which

$$\boldsymbol{\gamma}^0\boldsymbol{\gamma}^0 = \mathbf{I}, \quad \boldsymbol{\gamma}^0\boldsymbol{\gamma}^1 + \boldsymbol{\gamma}^1\boldsymbol{\gamma}^0 = \mathbf{0}, \quad \boldsymbol{\gamma}^1\boldsymbol{\gamma}^1 = -\mathbf{I} \tag{3}$$

supply a more explicit account. The Klein-Gordon operator can then be rendered

$$\Box + \varkappa^2 = (\boldsymbol{\gamma}^\mu \partial_\mu - i\varkappa)(\boldsymbol{\gamma}^\mu \partial_\mu + i\varkappa) \tag{4}$$

and we are led to the Dirac equation

$$(\boldsymbol{\gamma}^\mu \partial_\mu + i\varkappa)\,\psi = 0 \tag{5}$$

The objects $\boldsymbol{\gamma}^\mu$ cannot, by (3), be real/complex numbers. A pair of 2×2 *matrices* that do the trick are

$$\boldsymbol{\gamma}^0 \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \boldsymbol{\gamma}^1 \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{6}$$

Notice that while $\boldsymbol{\gamma}^0$ is hermitian, $\boldsymbol{\gamma}^1$ is antihermitian. The hermitian matrix

$$\boldsymbol{G} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \text{ which, as it happens, is just } \boldsymbol{\gamma}^0 \tag{7}$$

is, however, available as a "hermitianizer" in the sense[1] that it renders $\boldsymbol{G}\boldsymbol{\gamma}^0$ and $\boldsymbol{G}\boldsymbol{\gamma}^1$ simultaneously hermitian.

Looking back now to (5) we see that ψ has become a 2-component wave function, and that were we wrote $i\varkappa$ we might more properly have written $i\varkappa\mathbf{I}$. The "toy Dirac equation" (5) is a coupled pair of equations, which can be spelled out

$$\begin{pmatrix} i\varkappa & \partial_0 - \partial_1 \\ \partial_0 + \partial_1 & i\varkappa \end{pmatrix} \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{8}$$

Multiplication by the conjugated operator yields two copies of the Klein-Gordon equation:

$$\begin{pmatrix} -i\varkappa & \partial_0 - \partial_1 \\ \partial_0 + \partial_1 & -i\varkappa \end{pmatrix} \begin{pmatrix} i\varkappa & \partial_0 - \partial_1 \\ \partial_0 + \partial_1 & i\varkappa \end{pmatrix} = \begin{pmatrix} \Box + \varkappa^2 & 0 \\ 0 & \Box + \varkappa^2 \end{pmatrix}$$

Lagrangian formulation. It is (see CFT, Chapter 2, p. 24) the existence of a hermitianizer that provides access to the methods of Lagrangian field theory. We introduce

$$\begin{aligned} \mathfrak{L} &= -\hbar c\Big[i\tfrac{1}{2}\big\{\tilde{\psi}_\mu \boldsymbol{\gamma}^\mu \psi - \tilde{\psi}\,\boldsymbol{\gamma}^\mu \psi_\mu\big\} + \varkappa\,\tilde{\psi}\psi\Big] \\ &= \hbar c\left[\frac{\tilde{\psi}_\mu \boldsymbol{\gamma}^\mu \psi - \tilde{\psi}\,\boldsymbol{\gamma}^\mu \psi_\mu}{2i} - \varkappa\,\tilde{\psi}\psi\right] \end{aligned} \tag{9.0}$$

with $\tilde{\psi} \equiv \psi^\dagger \boldsymbol{G}$ and $\psi_\mu \equiv \partial_\mu \psi$; the $\hbar c$-factor has been introduced in order to ensure that $[\mathfrak{L}] =$ energy/length, and in the presumption that $[\tilde{\psi}\psi] = 1/\text{length}$,

[1] Compare (2–55) in CLASSICAL FIELD THEORY (1999). I will henceforth write CFT when referring to that source.

while the minus sign is physically inconsequential/cosmetic. Gauge-equivalent to $\mathcal{L}$ are

$$\mathcal{L}_1 \equiv \mathcal{L} + \tfrac{1}{2}i\hbar c\partial_\mu(\tilde{\psi}\boldsymbol{\gamma}^\mu\psi) = +\hbar c\big[\tilde{\psi}(i\boldsymbol{\gamma}^\mu\psi_\mu - \varkappa\psi)\big] \tag{9.1}$$

$$\mathcal{L}_2 \equiv \mathcal{L} - \tfrac{1}{2}i\hbar c\partial_\mu(\tilde{\psi}\boldsymbol{\gamma}^\mu\psi) = -\hbar c\big[(\tilde{\psi}_\mu\boldsymbol{\gamma}^\mu i + \tilde{\psi}\varkappa)\psi\big] \tag{9.2}$$

which permit one to write

$$\mathcal{L} = \tfrac{1}{2}(\mathcal{L}_1 + \mathcal{L}_2)$$

The $\tilde{\psi}^{\text{th}}$ Lagrange derivative operator looks most simply to (9.1), and yields the Dirac equation (5) as a field equation. The ψ^{th} Lagrange derivative of (9.2) yields as the other field equation $\tilde{\psi}_\mu\boldsymbol{\gamma}^\mu i + \tilde{\psi}\varkappa = 0$, which is seen by the following little argument

$$\begin{aligned}(\tilde{\psi}_\mu\boldsymbol{\gamma}^\mu i + \tilde{\psi}\varkappa) &= i(\tilde{\psi}_\mu\boldsymbol{\gamma}^\mu - i\tilde{\psi}\varkappa)\\ &= i\,(\boldsymbol{\gamma}^\mu\psi_\mu + i\varkappa\,\psi)^\dagger\boldsymbol{G}\end{aligned}$$

to be in effect the conjugate transpose of (5).

From the simple design of (9.1) it follows that if ψ is a *solution* of the Dirac equation (5) then $\mathcal{L}_1 = 0$, and by quick extension of that argument we learn that

$$\mathcal{L},\ \mathcal{L}_1 \text{ and } \mathcal{L}_2 \text{ vanish numerically if } \psi \text{ satisfies the Dirac equation} \tag{10}$$

From the manifest reality of $\mathcal{L}$ it follows by Noether's theorem[2] that

$$\partial_\mu Q^\mu = 0 \quad \text{where} \quad Q^\mu \equiv c\tilde{\psi}\boldsymbol{\gamma}^\mu\psi \tag{11}$$

where the factor c has been introduced so as to achieve $[Q^\mu] = 1/\text{time}$ (which in 2-dimensional spacetime is the dimension of "number flux"). Direct verification of (11) is easily accomplished:

$$\begin{aligned}\partial_\mu Q^\mu &= c\tilde{\psi}_\mu\boldsymbol{\gamma}^\mu\psi + c\tilde{\psi}\boldsymbol{\gamma}^\mu\psi_\mu\\ &= c(\tilde{\psi}i\varkappa)\psi + c\tilde{\psi}(-i\varkappa\,\psi) \quad \text{by the field equations}\\ &= 0\end{aligned}$$

Recalling the definitions (6) and (7), we have these explicit formulae:

$$\left.\begin{aligned}Q^0 &= \psi^{*0}\psi^0 + \psi^{*1}\psi^1\\ Q^1 &= \psi^{*0}\psi^0 - \psi^{*1}\psi^1\end{aligned}\right\} \tag{12}$$

Noether has supplied this generic description of the stress-energy tensor:[3]

$$S^\mu{}_\nu = \tilde{\psi}_\nu\frac{\partial\mathcal{L}}{\partial\tilde{\psi}_\mu} + \frac{\partial\mathcal{L}}{\partial\psi_\mu}\,\psi_\nu - \mathcal{L}\delta^\mu{}_\nu \quad \uparrow\text{drop, in consequence of (10)}$$

[2] See CFT (3–69).

[3] See CFT (1–34).

Working from (9.0), we obtain

$$S^{\mu}{}_{\nu} = \hbar c \left[\frac{\tilde{\psi}_\nu \boldsymbol{\gamma}^\mu \psi - \tilde{\psi} \boldsymbol{\gamma}^\mu \psi_\nu}{2i} \right] \tag{13}$$

in connection with which we observe that

$$\begin{aligned} \partial_\mu S^{\mu}{}_{\nu} &\sim \tilde{\psi}_{\mu\nu} \boldsymbol{\gamma}^\mu \psi + \tilde{\psi}_\nu \boldsymbol{\gamma}^\mu \psi_\mu - \tilde{\psi}_\mu \boldsymbol{\gamma}^\mu \psi_\nu - \tilde{\psi} \boldsymbol{\gamma}^\mu \psi_{\mu\nu} \\ &= (\tilde{\psi} i\boldsymbol{\varkappa})_\nu \psi + \tilde{\psi}_\nu (-i\boldsymbol{\varkappa}\, \psi) - (\tilde{\psi} i\boldsymbol{\varkappa}) \psi_\nu - \tilde{\psi} (-i\boldsymbol{\varkappa}\, \psi)_\nu \text{ by field equations} \\ &= 0 \end{aligned}$$

Notice that $S_{\mu\nu}$ is *not symmetric*: in this respect also toy Dirac theory is found to mimic precisely its 4-dimensional prototype.[4] Had we worked from (9.1) or (9.2) we would have been led to distinct but similar results.

Lorentz covariance. In 2-dimensional spacetime we write

$$\Lambda: \quad x \longrightarrow X = \boldsymbol{\Lambda} x \quad \text{with} \quad \boldsymbol{\Lambda}^{\mathsf{T}} \boldsymbol{g} \boldsymbol{\Lambda} = \boldsymbol{g} \tag{14}$$

to describe a Lorentz transformation. Necessarily $\det \boldsymbol{\Lambda} = \pm 1$. Infinitesimal Lorentz transformations are necessarily proper, and can be described

$$\boldsymbol{\Lambda} = \mathsf{I} + \boldsymbol{\alpha} + \cdots \quad \text{with} \quad \boldsymbol{\alpha}^{\mathsf{T}} \boldsymbol{g} + \boldsymbol{g} \boldsymbol{\alpha} = \mathbf{0} \;\Rightarrow\; \boldsymbol{\alpha} = \delta\omega \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{15}$$

By iteration we obtain

$$\boldsymbol{\Lambda} = \exp \left\{ \omega \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right\} = \begin{pmatrix} \cosh\omega & \sinh\omega \\ \sinh\omega & \cosh\omega \end{pmatrix} \tag{16}$$

where $\beta = \tanh\omega$ serves (via $\cosh\omega = 1/\sqrt{1-\beta^2} \equiv \gamma$ and $\sinh\omega = \beta/\gamma$) to establish explicit contact with kinematic aspects of the theory of Lorentz transformations. We observe that in 2-dimensional spacetime all (proper) Lorentz transformations have the irrotational character of "boosts."

The components of multicomponent fields are assumed to fold among themselves in linear representation of the Lorentz group $O(1,1)$:[5]

$$\Lambda: \quad \psi \longrightarrow \Psi = \boldsymbol{U}(\Lambda) \psi \tag{17.1}$$

First partials of such a field therefore transform

$$\Lambda: \quad \psi_\mu \longrightarrow \Psi_\mu = \boldsymbol{U}(\Lambda) \Lambda^{\mu}{}_{\nu} \psi_\mu \tag{17.2}$$

[4] Compare CFT (2–99).

[5] For omitted details see CFT, Chapter 2, pp. 32 *et seq*.

The Lorentz invariance of the Dirac Lagrangian (9.0) can be shown to entail (compare CFT (2–70))

$$\boldsymbol{U}^{-1} = \boldsymbol{G}^{-1}\boldsymbol{U}^{\dagger}\boldsymbol{G} \tag{18.1}$$

$$\boldsymbol{U}^{-1}\boldsymbol{\gamma}^{\mu}\boldsymbol{U} = \Lambda^{\mu}{}_{\nu}\boldsymbol{\gamma}^{\nu} \tag{18.2}$$

The first of those conditions asserts the "$\boldsymbol{G}$-unitarity" of the 2×2 matrix $\boldsymbol{U}$, and would reduce to ordinary unitarity if it were the case that $\boldsymbol{G} = \mathsf{I}$. Write

$$\boldsymbol{U} = \mathsf{I} + \boldsymbol{\beta} + \cdots$$

Then (18.1) entails the "$\boldsymbol{G}$-antihermiticity" of $\boldsymbol{\beta}$ (i.e., that $\boldsymbol{G}\boldsymbol{\beta} + (\boldsymbol{G}\boldsymbol{\beta})^{\dagger} = \mathbf{0}$) while (18.2) gives

$$\boldsymbol{\gamma}^{\mu}\boldsymbol{\beta} - \boldsymbol{\beta}\boldsymbol{\gamma}^{\mu} = \alpha^{\mu\nu}\boldsymbol{\gamma}_{\nu} \tag{19}$$

Stealing now from I know not whom, I claim that $\boldsymbol{\beta}$ can therefore be described

$$\boldsymbol{\beta} = \tfrac{1}{8}\alpha^{\rho\sigma}(\boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}_{\sigma} - \boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}_{\rho}) \tag{20}$$

and do so on the basis of the following argument: write $\boldsymbol{\gamma}^{\mu}(\boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}_{\sigma} - \boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}_{\rho})$ and use $\boldsymbol{\gamma}^{\mu}\boldsymbol{\gamma}^{\nu} + \boldsymbol{\gamma}^{\nu}\boldsymbol{\gamma}^{\mu} = 2g^{\mu\nu}\mathsf{I}$, in the form $\boldsymbol{\gamma}^{\mu}\boldsymbol{\gamma}_{\rho} = 2\delta^{\mu}{}_{\rho}\mathsf{I} - \boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}^{\mu}$, to pull the $\boldsymbol{\gamma}^{\mu}$ factors through to the right; one obtains

$$\begin{aligned}\boldsymbol{\gamma}^{\mu}(\boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}_{\sigma} - \boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}_{\rho}) &= 2(\delta^{\mu}{}_{\rho}\boldsymbol{\gamma}_{\sigma} - \delta^{\mu}{}_{\sigma}\boldsymbol{\gamma}_{\rho}) - \boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}^{\mu}\boldsymbol{\gamma}_{\sigma} + \boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}^{\mu}\boldsymbol{\gamma}_{\rho}\\ &= 2(\delta^{\mu}{}_{\rho}\boldsymbol{\gamma}_{\sigma} - \delta^{\mu}{}_{\sigma}\boldsymbol{\gamma}_{\rho}) - 2(\boldsymbol{\gamma}_{\rho}\delta^{\mu}{}_{\sigma} - \boldsymbol{\gamma}_{\sigma}\delta^{\mu}{}_{\rho})\\ &\qquad + \boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}^{\mu} - \boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}^{\mu}\end{aligned}$$

giving

$$\begin{aligned}\alpha^{\rho\sigma}\left[\boldsymbol{\gamma}^{\mu}(\boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}_{\sigma} - \boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}_{\rho}) - (\boldsymbol{\gamma}_{\rho}\boldsymbol{\gamma}_{\sigma} - \boldsymbol{\gamma}_{\sigma}\boldsymbol{\gamma}_{\rho})\boldsymbol{\gamma}^{\mu}\right] &= \alpha^{\rho\sigma}\left[4(\delta^{\mu}{}_{\rho}\boldsymbol{\gamma}_{\sigma} - \delta^{\mu}{}_{\sigma}\boldsymbol{\gamma}_{\rho})\right]\\ &= 4(\alpha^{\mu\nu} - \alpha^{\nu\mu})\,\boldsymbol{\gamma}_{\nu}\\ &= 8\alpha^{\mu\nu}\boldsymbol{\gamma}_{\nu}\end{aligned}$$

which completes the demonstration. Looking back again to the definitions (1), (6) and (15) of $\boldsymbol{g}$, $\boldsymbol{\gamma}^0$, $\boldsymbol{\gamma}^1$ and $\|\alpha^{\mu}{}_{\nu}\|$ we obtain

$$\boldsymbol{\gamma}_0 = g_{0\beta}\boldsymbol{\gamma}^{\beta} = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \boldsymbol{\gamma}_1 = g_{1\beta}\boldsymbol{\gamma}^{\beta} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$$

and

$$\|\alpha^{\mu\nu}\| = \|\alpha^{\mu}{}_{\beta}\|\|g^{\beta\nu}\| = \begin{pmatrix}0 & -\omega\\ \omega & 0\end{pmatrix}$$

which when brought to (20) give

$$\boldsymbol{\beta} = \tfrac{1}{2}\begin{pmatrix}\omega & 0\\ 0 & -\omega\end{pmatrix} \tag{21}$$

Calculation confirms that the $\boldsymbol{\beta}$ thus described does in fact satisfy (19). By trivial iteration we have

$$\boldsymbol{U} = \exp\left\{\tfrac{1}{2}\begin{pmatrix}\omega & 0\\ 0 & -\omega\end{pmatrix}\right\} = \begin{pmatrix}e^{+\frac{1}{2}\omega} & 0\\ 0 & e^{-\frac{1}{2}\omega}\end{pmatrix} \tag{22}$$

Calculation gives

$$\boldsymbol{U}^{-1}\boldsymbol{\gamma}^0\boldsymbol{U} = \begin{pmatrix}0 & e^{-\omega}\\ e^{+\omega} & 0\end{pmatrix} = \cosh\omega\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + \sinh\omega\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$$
$$\boldsymbol{U}^{-1}\boldsymbol{\gamma}^1\boldsymbol{U} = \begin{pmatrix}0 & -e^{-\omega}\\ e^{+\omega} & 0\end{pmatrix} = \sinh\omega\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + \cosh\omega\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$$

in exact agreement with (18.2).

Let us look now more closely to the implications of (18.1). Necessarily

$$\det\boldsymbol{U} = e^{i\phi} \tag{23.1}$$

so $\boldsymbol{G}$-unitary matrices can (in the 2-dimensional case) be displayed

$$\boldsymbol{U} = e^{\frac{1}{2}i\phi}\cdot\boldsymbol{S} \tag{23.2}$$

where $\boldsymbol{S}$ is a *unimodular* $\boldsymbol{G}$-unitary matrix:

$$\boldsymbol{S}^{-1} = \boldsymbol{G}^{-1}\boldsymbol{S}^\dagger\boldsymbol{G} \quad \text{with} \quad \det\boldsymbol{S} = 1 \tag{23.3}$$

The most general such 2×2 matrix is found to have the form

$$\boldsymbol{S} = \begin{pmatrix}a & ib\\ ic & d\end{pmatrix} \quad \text{with} \quad a,b,c,d \text{ real and } ad+bc=1 \tag{23.4}$$

If we write $\boldsymbol{S} = e^{i\boldsymbol{H}}$ then $\boldsymbol{S}$ will be unimodular if and only if $\mathrm{tr}\boldsymbol{H} = 0$, and $\boldsymbol{G}$-unitary if and only if $\boldsymbol{H}$ is $\boldsymbol{G}$-hermitian

$$\boldsymbol{H} = \boldsymbol{G}^{-1}\boldsymbol{H}^\dagger\boldsymbol{G} \tag{23.5}$$

which requires that $\boldsymbol{GH}$ be hermitian in the standard sense; i.e., that

$$\boldsymbol{H} = \boldsymbol{G}^{-1}\cdot\begin{pmatrix}p & r+is\\ r-is & q\end{pmatrix} = \begin{pmatrix}r-is & q\\ p & r+is\end{pmatrix} \tag{24.6}$$

which will be traceless if and only if $r=0$. Looking in the light of these formal developments back to (22), we see that the $\boldsymbol{U}$ encountered there is unimodular—an instance of (23.4) with $b=c=0$ and $d=a^{-1}$.

We have arrived at the association

$$\boldsymbol{\Lambda}(\omega) = \begin{pmatrix}\cosh\omega & \sinh\omega\\ \sinh\omega & \cosh\omega\end{pmatrix} \quad\Longleftrightarrow\quad \boldsymbol{U}(\omega) = \begin{pmatrix}e^{+\frac{1}{2}\omega} & 0\\ 0 & e^{-\frac{1}{2}\omega}\end{pmatrix} \tag{25}$$

from which it is already clear that if Lorentz covariance of the toy Dirac theory is to be achieved then the 2-component wave function ψ cannot transform as a 2-vector, but must transform by the distinctive rule

$$\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix} \longrightarrow \begin{pmatrix}\Psi^1\\ \Psi^2\end{pmatrix} = \boldsymbol{U}(\omega)\begin{pmatrix}\psi^1\\ \psi^2\end{pmatrix}$$

$$\boldsymbol{U}(\omega) = \begin{pmatrix}\cosh\frac{1}{2}\omega + \sinh\frac{1}{2}\omega & 0\\ 0 & \cosh\frac{1}{2}\omega - \sinh\frac{1}{2}\omega\end{pmatrix}$$

$$= \begin{pmatrix}\sqrt{\frac{1+\beta}{1-\beta}} & 0\\ 0 & \sqrt{\frac{1-\beta}{1+\beta}}\end{pmatrix}$$

Notice that $\boldsymbol{U}$ has turned out to be unimodular, though we nowhere had reason to insist that it be so.

The intrusion of (what we might informally call) half-angles speaks to us of the familiar double-valuedness of the spinor representations of $O(3)\ldots$ and indeed: it follows transparently from (18) that if $\boldsymbol{U}$ satifies that pair of conditions, then so also does its negative. So in place of (25) we might more properly write

$$\boldsymbol{\Lambda}(\omega) \quad \Longleftrightarrow \quad \pm\boldsymbol{U}(\omega) \tag{26}$$

In the present context, however, the two branches of the *spinor representation of the Lorentz group* $O(1,1)$ are—uncharacteristically—disjoint. They can be connected, but at cost of a *complexification of the "rapidity" parameter*:

$$\boldsymbol{\Lambda}(\omega + 2\pi i) = \boldsymbol{\Lambda}(\omega) \quad \text{but} \quad \boldsymbol{U}(\omega + 2\pi i) = -\boldsymbol{U}(\omega) \tag{27}$$

We noticed that only a 1-parameter subgroup of the full 3-parameter group $SU(2;\boldsymbol{G})$ of unimodular $\boldsymbol{G}$-unitary matrices is pressed into service to represent the action of Lorentzian boosts, and inquire now into the question of whether or not that subgroup is in any way "distinguished." For an arbitrary 2×2 matrix $\boldsymbol{M}$ one has $\det(\boldsymbol{M} - \lambda\boldsymbol{I}) = \lambda^2 - \lambda\cdot\mathrm{tr}\boldsymbol{M} + \det\boldsymbol{M}$, so if $\boldsymbol{M}$ is traceless then the Cayley-Hamilton theorem supplies $\boldsymbol{M}^2 + (\det\boldsymbol{M})\boldsymbol{I} = \boldsymbol{0}$, whence

$$e^{\boldsymbol{M}} = \cos\theta\cdot\boldsymbol{I} + \sin\theta\cdot\boldsymbol{M}/\theta \quad \text{with} \quad \theta \equiv \sqrt{\det\boldsymbol{M}}$$

In application of these general remarks we have

$$\boldsymbol{S} = e^{i\boldsymbol{H}} = \cos\theta\cdot\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \sin\theta\cdot\frac{1}{\theta}\begin{pmatrix}s & iq\\ ip & -s\end{pmatrix} \quad \text{with} \quad \theta \equiv \sqrt{pq - s^2}$$

which reproduces the design of (23.4). Boosts have been found to be associated with the case $p = q = 0$, where we have

$$= \cos is\cdot\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \sin is\cdot\frac{1}{is}\begin{pmatrix}s & 0\\ 0 & -s\end{pmatrix}$$

$$= \cosh s\cdot\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \sinh s\cdot\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

We recover (25) when, as stipulated at (21), we set $s = \frac{1}{2}\omega$. Let us now, at (24.6), set $p = u + v$ and $q = u - v$ so as to obtain

$$\boldsymbol{H}_{\text{traceless}} = \begin{pmatrix} -is & u-v \\ u+v & +is \end{pmatrix} \equiv s\boldsymbol{\tau}_0 + u\boldsymbol{\tau}_1 + v\boldsymbol{\tau}_2$$

Then

$$\det(s\boldsymbol{\tau}_0 + u\boldsymbol{\tau}_1 + v\boldsymbol{\tau}_2) = s^2 - u^2 + v^2$$

Each $\boldsymbol{\tau}$-matrix is traceless; their determinants are given by

$$\det\boldsymbol{\tau}_0 = +1; \quad \det\boldsymbol{\tau}_1 = -1; \quad \det\boldsymbol{\tau}_3 = +1$$

and their multiplicative properties present a kind of twisted version of the Pauli algebra:

$$\boldsymbol{\tau}_0^2 = -1; \quad \boldsymbol{\tau}_1^2 = +1; \quad \boldsymbol{\tau}_2^2 = -1$$

$$\begin{aligned} \boldsymbol{\tau}_0\boldsymbol{\tau}_1 &= -\boldsymbol{\tau}_1\boldsymbol{\tau}_0 = +i\boldsymbol{\tau}_2 \\ \boldsymbol{\tau}_1\boldsymbol{\tau}_2 &= -\boldsymbol{\tau}_2\boldsymbol{\tau}_1 = +i\boldsymbol{\tau}_0 \\ \boldsymbol{\tau}_2\boldsymbol{\tau}_0 &= -\boldsymbol{\tau}_0\boldsymbol{\tau}_2 = -i\boldsymbol{\tau}_1 \end{aligned}$$

If there is a "distinguished" element it would appear to be not $\boldsymbol{\tau}_0$ but $\boldsymbol{\tau}_1$. I had hoped to be able to assign meaning to the "spin" of the ψ-field, even though the toy theory does not support a concept of (orbital) angular momentum... but appear to be simply going in circles, so abandon this aspect of my topic. In the absence of a theory of spin it appears to be impossible to use Belinfante's trick (CFT, Chapter 2, p. 43) to achieve symmetrization of the stress-energy tensor.

Clifford algebras—especially the algebras of order 2. It appears to have been William Clifford who first undertook to take the "square root of a quadratic form," writing

$$p_1^2 + p_2^2 + \cdots + p_n^2 = (\boldsymbol{\epsilon}^1 p_1 + \boldsymbol{\epsilon}^2 p_2 + \cdots + \boldsymbol{\epsilon}^n p_n)^2 \tag{28}$$

and imposing upon the "hypernumbers" $\boldsymbol{\epsilon}^i$ the requirements that

$$\left.\begin{aligned} (\boldsymbol{\epsilon}^i)^2 = 1 \quad &: \quad i = 1, 2, \ldots, n \\ \boldsymbol{\epsilon}^i\boldsymbol{\epsilon}^j = -\boldsymbol{\epsilon}^j\boldsymbol{\epsilon}^i \quad &: \quad i \neq j \end{aligned}\right\} \tag{29}$$

though precisely why he did so—and where/when—remains unclear: persual of his *Collected Papers*[6] reveals an abiding interest in the geometric algebras of Hamilton and Grassmann, but it appears that what we now call "Clifford algebra" can be detected in Clifford's own work only as fragmentary hints (see the editor's remark at p. *lxvii* in the volume just cited), written near the end of his brief life (1845–1879). Pertti Lounestoe informs us[7] that Clifford's idea was

[6] The edition of 1882 was reprinted by Chelsea in 1968.

[7] *Clifford Algebras and Spinors* (1997), p. 9.

first mentioned in a talk (1876), the text of which was published posthumously as "XLIII. On the classification of geometrical algebras" in *Collected Papers*. Clifford's inspired murmuring attracted very little attention, however, until—fifty years later, and stripped of its geometrical overlay—the algebraic essence of the idea occurred independently to Dirac, and was put by him to stunningly productive work. The literature, some of it of intimidating complexity, has by become vast, and the theory of Clifford algebras (like quaternion algebra before it) has displayed a curious power to make fanatics of otherwise staid and sensible applied mathematicians and physicists. We will be concerned only with the simplest elements of the subject.[8]

Higher powers of the expression which appears squared on the right side of (28) are typified by the following:

$$\begin{aligned}&(\boldsymbol{\epsilon}^1 p_1 + \boldsymbol{\epsilon}^2 p_2 + \cdots + \boldsymbol{\epsilon}^n p_n)^7 \\ &\qquad = \text{linear combination of terms of the type } \boldsymbol{\epsilon}^{i_1}\boldsymbol{\epsilon}^{i_2}\boldsymbol{\epsilon}^{i_3}\boldsymbol{\epsilon}^{i_4}\boldsymbol{\epsilon}^{i_5}\boldsymbol{\epsilon}^{i_6}\boldsymbol{\epsilon}^{i_7}\end{aligned}$$

Look to a typical such term: drawing upon (28) we have (on the presumption that $n \geq 9$)

$$\begin{aligned}\boldsymbol{\epsilon}^5\boldsymbol{\epsilon}^2\boldsymbol{\epsilon}^1\boldsymbol{\epsilon}^1\boldsymbol{\epsilon}^7\boldsymbol{\epsilon}^2\boldsymbol{\epsilon}^1 &= -\boldsymbol{\epsilon}^1\boldsymbol{\epsilon}^1\boldsymbol{\epsilon}^1\boldsymbol{\epsilon}^2\boldsymbol{\epsilon}^2\boldsymbol{\epsilon}^5\boldsymbol{\epsilon}^7 \\ &\qquad \uparrow \text{odd number of permutations required to achieve ascending order} \\ &= -\boldsymbol{\epsilon}^1\boldsymbol{\epsilon}^5\boldsymbol{\epsilon}^7 \quad \text{because} \begin{cases} (\boldsymbol{\epsilon}^i)^{\text{even}} = \mathsf{I} \\ (\boldsymbol{\epsilon}^i)^{\text{odd}} = \boldsymbol{\epsilon}^i \end{cases}\end{aligned}$$

Evidently every $\boldsymbol{\epsilon}$-product encountered in such an expression can, by (28), be brought to the form of one or another of the products which appear in the following list:

$$(\mathsf{I} \text{ else } \boldsymbol{\epsilon}^1) \cdot (\mathsf{I} \text{ else } \boldsymbol{\epsilon}^2) \cdots (\mathsf{I} \text{ else } \boldsymbol{\epsilon}^n)$$

The list asks us to make n binary choices, and therefore presents a total of 2^n distinct "canonical products." We are led thus to contemplate expressions of the design

$$\begin{aligned}\boldsymbol{A} = a\mathsf{I} + \sum_i a_i \boldsymbol{\epsilon}^i + \sum_{i<j} a_{ij}\boldsymbol{\epsilon}^i\boldsymbol{\epsilon}^j + \cdots + \sum_{i_1<i_2<\cdots<i_p} a_{i_1 i_2 \cdots i_p}\boldsymbol{\epsilon}^{i_1}\boldsymbol{\epsilon}^{i_2}\cdots\boldsymbol{\epsilon}^{i_p} \\ + a_{12\cdots n}\boldsymbol{\epsilon}^1\boldsymbol{\epsilon}^2\cdots\boldsymbol{\epsilon}^n\end{aligned} \tag{30}$$

where the coefficients are taken to be (let us say) real numbers, and where there are evidently $\binom{n}{p}$ terms of order p. The set of all such "Clifford numbers" $\boldsymbol{A}$ is closed under both addition and multiplication, and is called the *Clifford algebra* $\mathcal{C}\ell_n$, of which $\{\boldsymbol{\epsilon}^1, \boldsymbol{\epsilon}^2, \ldots, \boldsymbol{\epsilon}^n\}$ are the "generators."

[8] For good accounts see (for example) §§3–5 in P. K. Raševskiĭ, "The theory of spinors," Amarican Mathematical Society Translations, Series 2, Volume 6 (1957) or the recent text by P. Lounestoe (cited above).

To place $\mathcal{C}\ell_n$ in its larger context: an *associative linear algebra* $\mathfrak{A}$ is a vector space—let the elements be notated

$$\boldsymbol{A} = a^1\boldsymbol{e}_1 + a^2\boldsymbol{e}_2 + \cdots + a^m\boldsymbol{e}_m \equiv a^i\,\boldsymbol{e}_i$$

—on which a law of multiplication is defined

$$\begin{aligned}\boldsymbol{AB} &= a^i b^j\,\boldsymbol{e}_i\boldsymbol{e}_j \\ \boldsymbol{e}_i\boldsymbol{e}_j &\equiv \textstyle\sum_p c_i{}^p{}_j\,\boldsymbol{e}_p\end{aligned} \tag{31.1}$$

and is required, moreover, to be associative:

$$\boldsymbol{A}(\boldsymbol{BC}) = (\boldsymbol{AB})\boldsymbol{C} \quad\Longleftrightarrow\quad c_i{}^q{}_p\,c_j{}^p{}_k = c_i{}^p{}_j\,c_p{}^q{}_k$$

Commutivity ($c_i{}^p{}_j = c_j{}^p{}_i$) is, however, typically *not* required. If we define $\mathbb{E}_i \equiv \|c_i{}^q{}_p\|$ then the associativity condition can be expressed

$$\mathbb{E}_i\,\mathbb{E}_j = \textstyle\sum_p c_i{}^p{}_j\,\mathbb{E}_p \tag{31.2}$$

which shows that every associative algebra of order m admits of $m \times m$ matrix representation (this is the non-trivial converse of the trivial statement that every matrix algebra is associative), but leaves open the question of whether $\mathfrak{A}$ admits of lower-dimensional matrix representation. Notice that (31.1) is—insofar as it involves weighted summation—more general than would be permitted by stipulation that the $\boldsymbol{e}_i$ are elements of a finite group, though every such group provides a specialized instance of (31.1).

It is by now apparent that $\mathcal{C}\ell_n$ is a specialized associative linear algebra of order $m = 2^n$, and admits of $2^n \times 2^n$ matrix representation. It is, however, not presently clear how to construct the least-dimensional representation of $\mathcal{C}\ell_n$. Or how to construct $\boldsymbol{A}^{-1}$, or even how to decide (directly, without recourse to matrix representation theory) whether $\boldsymbol{A}^{-1}$ exists. These are typical of issues taken up in the literature.

Relativity—and a host of other pure/applied topics as well—inspire interest also in *indefinite* quadratic forms

$$p_1^2 + p_2^2 + \cdots + p_n^2 - q_1^2 - q_2^2 - \cdots - q_m^2$$

The associated Clifford algebras—which arise from writing

$$(\boldsymbol{\epsilon}^1 p_1 + \cdots + \boldsymbol{\epsilon}^n p_n + \boldsymbol{\epsilon}^{n+1} q_1 + \cdots + \boldsymbol{\epsilon}^{n+m} q_m)^2$$

and requiring that

$$\left.\begin{aligned}(\boldsymbol{\epsilon}^i)^2 &= \begin{cases} +\mathsf{I} & \text{if} \quad i = 1, 2, \ldots, n \\ -\mathsf{I} & \text{if} \quad i = n+1, n+2, \ldots, m \end{cases} \\ \boldsymbol{\epsilon}^i\,\boldsymbol{\epsilon}^j &= -\boldsymbol{\epsilon}^j\,\boldsymbol{\epsilon}^i \quad \text{if} \quad i \neq j\end{aligned}\right\} \tag{32}$$

—are denoted $\mathcal{C}\ell_{n,m}$. The algebra previously designated $\mathcal{C}\ell_n$ would in this refined notation be designated $\mathcal{C}\ell_{n,0}$.

Evidently $\mathcal{C}\ell_{0,1}$ is just $\mathbb{C}$, the algebra of complex numbers, which we are in position now to observe admits of real 2×2 matrix representation. Look to the details: (32) reduces to the statement

$$\boldsymbol{\epsilon}^2 = -\mathbf{I}$$

Write $\mathbf{I} \to \boldsymbol{e}_1$ and $\boldsymbol{\epsilon} \to \boldsymbol{e}_2$ to establish contact with the generic language of (31.1). Then $\mathbf{I}\cdot\mathbf{I} = \mathbf{I}, \mathbf{I}\cdot\boldsymbol{\epsilon} = \boldsymbol{\epsilon}, \boldsymbol{\epsilon}\cdot\mathbf{I} = \boldsymbol{\epsilon}, \boldsymbol{\epsilon}\cdot\boldsymbol{\epsilon} = -\mathbf{I}$ read

$$\begin{aligned}
\boldsymbol{e}_1\boldsymbol{e}_1 &= c_1{}^1{}_1\boldsymbol{e}_1 + c_1{}^2{}_1\boldsymbol{e}_2 = \quad 1\boldsymbol{e}_1 + 0\boldsymbol{e}_2 \\
\boldsymbol{e}_1\boldsymbol{e}_2 &= c_1{}^1{}_2\boldsymbol{e}_1 + c_1{}^2{}_2\boldsymbol{e}_2 = \quad 0\boldsymbol{e}_1 + 1\boldsymbol{e}_2 \\
\boldsymbol{e}_2\boldsymbol{e}_1 &= c_2{}^1{}_1\boldsymbol{e}_1 + c_2{}^2{}_1\boldsymbol{e}_2 = \quad 0\boldsymbol{e}_1 + 1\boldsymbol{e}_2 \\
\boldsymbol{e}_2\boldsymbol{e}_2 &= c_2{}^1{}_2\boldsymbol{e}_1 + c_2{}^2{}_2\boldsymbol{e}_2 = -1\boldsymbol{e}_1 + 0\boldsymbol{e}_2
\end{aligned}$$

giving

$$\mathbb{E}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \to \mathbb{I} \qquad \text{and} \qquad \mathbb{E}_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \to \mathbb{J}$$

The elements $\boldsymbol{z} = x\mathbf{I} + y\boldsymbol{\epsilon}$ of $\mathcal{C}\ell_{0,1}$ acquire therefore the matrix representations

$$\mathbb{Z} = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$$

Evidently $\det\mathbb{Z} = x^2 + y^2$ and $\mathbb{Z}^{-1} \to (x^2+y^2)^{-1}(x\mathbf{I} - y\boldsymbol{\epsilon}) = (\boldsymbol{z}^*\boldsymbol{z})^{-1}\boldsymbol{z}^*$ with $\boldsymbol{z}^* = x\mathbf{I} - y\boldsymbol{\epsilon}$. I will, for methodological reasons, conclude with indication of how $\boldsymbol{z}^{-1}$ might have been obtained "directly, without recourse to matrix representation theory." Assume without significant loss of generality that $x = 1$ and write

$$\begin{aligned}
(\mathbf{I} + y\boldsymbol{\epsilon})^{-1} &= \mathbf{I} - y\boldsymbol{\epsilon} + (y\boldsymbol{\epsilon})^2 - (y\boldsymbol{\epsilon})^3 + (y\boldsymbol{\epsilon})^4 - (y\boldsymbol{\epsilon})^5 + \cdots \\
&= (1 - y^2 + y^4 - \cdots)\mathbf{I} - (y - y^3 + y^5 - \cdots)\boldsymbol{\epsilon} \\
&= (1+y^2)^{-1}(\mathbf{I} - y\boldsymbol{\epsilon})
\end{aligned}$$

The argument fails only if $x = 0$, but in that case one has

$$(y\boldsymbol{\epsilon})^{-1} = (y^2)^{-1}(-y\boldsymbol{\epsilon})$$

by inspection.

The algebra $\mathcal{C}\ell_{0,2}$ is familiar as the algebra $\mathbb{Q}$ of real quaternions, as I now demonstrate. The general element can be written

$$\begin{aligned}
\boldsymbol{q} &= w\mathbf{I} + x\boldsymbol{\epsilon}_1 + y\boldsymbol{\epsilon}_2 + z\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_2 \\
&= w\mathbf{I} + x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k} \quad \text{in Hamilton's notation}
\end{aligned}$$

and the conditions (32) become

$$\boldsymbol{i}^2 = \boldsymbol{j}^2 = \boldsymbol{k}^2 = -\mathsf{I} \tag{33.1}$$

$$\begin{aligned} \boldsymbol{i}\boldsymbol{j} = \boldsymbol{k} &= -\boldsymbol{j}\boldsymbol{i} \\ \boldsymbol{j}\boldsymbol{k} = \boldsymbol{i} &= -\boldsymbol{k}\boldsymbol{j} \\ \boldsymbol{k}\boldsymbol{i} = \boldsymbol{j} &= -\boldsymbol{i}\boldsymbol{k} \end{aligned} \tag{33.2}$$

of which

$$\begin{aligned} (a^1\boldsymbol{i} + a^2\boldsymbol{j} + a^3\boldsymbol{k})(b^1\boldsymbol{i} + b^2\boldsymbol{j} + b^3\boldsymbol{k}) = &-(a^1b^1 + a^2b^2 + a^3b^3)\mathsf{I} \\ &+ (a^2b^3 - a^3b^2)\boldsymbol{i} \\ &+ (a^3b^1 - a^1b^3)\boldsymbol{j} \\ &+ (a^1b^2 - a^2b^1)\boldsymbol{k} \end{aligned}$$

is a celebrated particular consequence. We are assured that $\mathcal{C}\ell_{0,2}$ admits of real 4×4 matrix representation, but I skip the detailed demonstration in order to confront the force of the question: How, in the absence of an explicit matrix representation, does one construct $\boldsymbol{q}^{-1}$? Look to the series expansion of

$$[\mathsf{I} + (x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})]^{-1}$$

and, using $(x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})^2 = -(x^2 + y^2 + z^2)\mathsf{I}$ (established just above), obtain

$$[\mathsf{I} + (x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})]^{-1} = (1 + x^2 + y^2 + z^2)^{-1}[\mathsf{I} - (x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})]$$

Quick tinkering (send $x \to x/w, y \to y/w, z \to z/w$ and simplify) leads from here to the conclusion (valid even when $w = 0$) that $\boldsymbol{q}^{-1}$ exists if and only if $\boldsymbol{q}^*\boldsymbol{q} = w^2 + x^2 + y^2 + z^2 \neq 0$ (i.e., for all $\boldsymbol{q} \neq \boldsymbol{0}$) and is given then by

$$\boldsymbol{q}^{-1} = (\boldsymbol{q}^*\boldsymbol{q})^{-1}\boldsymbol{q}^* \quad \text{with} \quad \boldsymbol{q}^* \equiv w\mathsf{I} - x\boldsymbol{i} - y\boldsymbol{j} - z\boldsymbol{k}$$

A moment's further tinkering leads—not at all to our surprise—to this complex 2×2 representation of the quaterion algebra:

$$\boldsymbol{q} \to \mathbb{Q} = \begin{pmatrix} w + z & -x - iy \\ x - iy & w - z \end{pmatrix} \equiv w\mathbb{U} + x\mathbb{I} + y\mathbb{J} + z\mathbb{K}$$

We have

$$\det \mathbb{Q} = w^2 + x^2 + y^2 + z^2$$

and observe that the matrices $\mathbb{I}$, $\mathbb{J}$ and $\mathbb{K}$ are traceless antihermitian. We recover the previously described representation of $\mathcal{C}\ell_{0,1}$ when $y = z = 0$.

$\mathcal{C}\ell_{1,3}$ is just the Dirac algebra—an algebra of order 16, generated by Dirac's celebrated $\boldsymbol{\gamma}^\mu$-matrices. It possesses a real 16×16 representation, but is well known to admit also of complex 4×4 representation.

Which brings us to $\mathcal{C}\ell_{1,1}$ —the "toy Dirac algebra" implicit in (1), of which we have been acquainted with the generators since (6) but have not previously had reason to regard as a closed algebra, with general element given by

$$\begin{aligned} \boldsymbol{d} = s\,\mathsf{I} + v_1\boldsymbol{\gamma}^1 + v_2\boldsymbol{\gamma}^2 + p\boldsymbol{\Gamma} \\ \boldsymbol{\Gamma} \equiv \boldsymbol{\gamma}^1\boldsymbol{\gamma}^2 \end{aligned} \tag{34}$$

$\mathcal{C}\ell_{1,1}$ is evidently a variant of the quaternion algebra; in place of (33) we have

$$(\boldsymbol{\gamma}^1)^2 = +\mathsf{I}\,, \quad (\boldsymbol{\gamma}^2)^2 = -\mathsf{I}\,, \quad (\boldsymbol{\Gamma})^2 = +\mathsf{I} \tag{35.1}$$

$$\begin{aligned} \boldsymbol{\gamma}^1\boldsymbol{\gamma}^2 &= \boldsymbol{\Gamma} \\ \boldsymbol{\gamma}^1\boldsymbol{\Gamma} &= \boldsymbol{\gamma}^2 \\ \boldsymbol{\gamma}^2\boldsymbol{\Gamma} &= \boldsymbol{\gamma}^1 \end{aligned} \tag{35.2}$$

A real 4×4 representation is assured, but a real (!) 2×2 representation is implicit in (6)

$$\boldsymbol{d} \to \mathbb{D} = \begin{pmatrix} s+p & v_1 - v_2 \\ v_1 + v_2 & s-p \end{pmatrix} \equiv s\,\mathbb{I} + v_1\,\mathbb{\Gamma}^1 + v_2\,\mathbb{\Gamma}^2 + p\,\mathbb{\Gamma} \tag{36}$$

and is much easier to work with. Define

$$\mathbb{D}^* = s\,\mathbb{I} - v_1\,\mathbb{\Gamma}^1 - v_2\,\mathbb{\Gamma}^2 - p\,\mathbb{\Gamma}$$

and obtain

$$\mathbb{D}^*\mathbb{D} = \underbrace{[\,s^2 - (v_1^2 - v_2^2) - p^2\,]}_{\det\mathbb{D}}\cdot\mathbb{I}$$

giving $\boldsymbol{d}^{-1} = (\boldsymbol{d}^*\boldsymbol{d})^{-1}\boldsymbol{d}^*$; this result is structurally quite familiar, but novel in respect to its detailed meaning.

The notations $s/v/p$ are intended to suggest *scalar/vector/pseudoscalar*, for reasons which I now discuss. Let

$$\boldsymbol{w} = a_1\boldsymbol{\gamma}^1 + a_2\boldsymbol{\gamma}^2 + b\boldsymbol{\Gamma} \quad \text{with} \quad a_1^2 - a_2^2 + b^2 = 1$$

Then $\boldsymbol{w}^2 = \mathsf{I}$ and

$$\boldsymbol{u} \equiv e^{\phi\boldsymbol{w}} = \cosh\phi\cdot\mathsf{I} + \sinh\phi\cdot\boldsymbol{w}$$

We look to the similarity transformation $\boldsymbol{d} \to \boldsymbol{D} = \boldsymbol{u}^{-1}\boldsymbol{d}\,\boldsymbol{u}$; i.e., to

$$\boldsymbol{D} = (\cosh\phi\cdot\mathsf{I} - \sinh\phi\cdot\boldsymbol{w})(s\mathsf{I} + v_1\boldsymbol{\gamma}^1 + v_2\boldsymbol{\gamma}^2 + p\boldsymbol{\Gamma})(\cosh\phi\cdot\mathsf{I} + \sinh\phi\cdot\boldsymbol{w})$$

which for infinitesimal ϕ becomes

$$\boldsymbol{d} \to \boldsymbol{D} = \boldsymbol{d} + \delta\phi\cdot[\boldsymbol{d}\,,\boldsymbol{w}] + \cdots$$

with

$$\begin{aligned}\delta\phi\cdot[\boldsymbol{d},\boldsymbol{w}] &= 2\delta\phi(bv_2-a_2p)\,\boldsymbol{\gamma}^1+2\delta\phi(bv_1-a_1p)\,\boldsymbol{\gamma}^2+2\delta\phi(a_2v_1-a_1v_2)\,\boldsymbol{\Gamma}\\ &= \delta v_1\boldsymbol{\gamma}^1+\delta v_2\boldsymbol{\gamma}^2+\delta p\,\boldsymbol{\Gamma}\end{aligned}$$

To exclude p-dependence from δv we find ourselves obliged to set $a_1=a_2=0$. Then $b=\pm1$; without loss of generality we set $b=1$, and obtain

$$\begin{aligned}\boldsymbol{D} &= e^{-\phi\boldsymbol{\Gamma}}\boldsymbol{d}\,e^{\phi\boldsymbol{\Gamma}}\\ &\qquad e^{\phi\boldsymbol{\Gamma}}=\cosh\phi\cdot\mathbf{I}+\sinh\phi\cdot\boldsymbol{\Gamma}\\ &= s\,\mathbf{I}+(v_1\cosh2\phi+v_2\sinh\phi)\,\boldsymbol{\gamma}^1+(v_1\sinh2\phi+v_2\cosh\phi)\,\boldsymbol{\gamma}^2+p\boldsymbol{\Gamma}\\ &= S\,\mathbf{I}+V_1\boldsymbol{\gamma}^1+V_2\boldsymbol{\gamma}^2+P\boldsymbol{\Gamma}\end{aligned}$$

We have here recovered the 2-dimensional Lorentz transformations as natural objects (and have at the same time indicated why it makes sense to use the term "vector" in reference to the v terms), but have left unexplained why p has been called a "pseudoscalar." Suppose, however, we were to set

$$\boldsymbol{u}=\boldsymbol{\gamma}^1\to\begin{pmatrix}0&1\\1&0\end{pmatrix}\quad:\quad\text{in this case }\boldsymbol{u}^{-1}=\boldsymbol{u}$$

The unimodular (or $\boldsymbol{d}^*\boldsymbol{d}$-preserving) transformation $\boldsymbol{d}\to\boldsymbol{D}=\boldsymbol{u}^{-1}\boldsymbol{d}\,\boldsymbol{u}$ is then improper (disjoint from the identity), and we compute

$$\boldsymbol{D}=s\,\mathbf{I}+v_1\boldsymbol{\gamma}^1-v_2\boldsymbol{\gamma}^2-p\boldsymbol{\Gamma}$$

The Lorentz transformation $(v_1,v_2)\to(v_1,-v_2)$ is improper in the familiar sense, and has sent $p\to-p$.

Clifford algebra when the metric is non-diagonal. Familiar transformations serve to achieve

$$\begin{pmatrix}g_{11}&g_{12}&\cdots&g_{1N}\\g_{21}&g_{22}&\cdots&g_{2N}\\\vdots&\vdots&\ddots&\vdots\\g_{N1}&g_{N2}&\cdots&g_{NN}\end{pmatrix}\to\begin{pmatrix}g_1&&&\\&g_2&&\\&&\ddots&\\&&&g_N\end{pmatrix}\to\begin{pmatrix}\pm1&&&\\&\pm1&&\\&&\ddots&\\&&&\pm1\end{pmatrix}$$

where the final matrix has n plus signs on the diagonal, and $m=N-n$ minus signs. The theory of $\mathcal{C}\ell_{n,m}$ standardly assumes such preparations to have been carried out. I have (general relativistic) interest, however, in seeing how one might contrive to work with a non-specialized real symmetric metric. The algebra which arises from writing $g^{ij}p_ip_j=(\boldsymbol{\epsilon}^ip_i)^2$ or, as I find now more convenient,

$$g_{ij}x^ix^j=(x^i\boldsymbol{\epsilon}_i)^2 \tag{37}$$

will be designated $\mathcal{C}\ell_N(g)$, and springs from anticommutation relations

$$\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_j+\boldsymbol{\epsilon}_j\boldsymbol{\epsilon}_i=2g_{ij} \tag{38}$$

of which (29) no longer provide an accurate description. The only surviving allusion to the possibility that the metric might be "indefinite" will reside in the observation that

$$g \equiv \det \|g_{ij}\| = (-)^m |g| \tag{39}$$

The important role formerly played by the antisymmetry condition (29) will be taken over by the statement that, in consequence of (38),

$$\boldsymbol{\epsilon}_{ij} \equiv \tfrac{1}{2}(\boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_j - \boldsymbol{\epsilon}_j\boldsymbol{\epsilon}_i) = \boldsymbol{\epsilon}_i\boldsymbol{\epsilon}_j - g_{ij}\,\mathbf{I} \quad \text{is antisymmetric} \tag{40}$$

In 1960 I had reason to work out the detailed theory of $\mathcal{C}\ell_4(g)$. Here, taking that work[9] as my model, I will look only to the simplest case: $\mathcal{C}\ell_2(g)$.

The general element of $\mathcal{C}\ell_2(g)$ can be written

$$\boldsymbol{G} = S\mathbf{I} + V^i\boldsymbol{\epsilon}_i + \tfrac{1}{2}T^{ij}\boldsymbol{\epsilon}_{ij} \tag{41}$$

where T^{ij} can without loss of generality be assumed to be antisymmetric; since 2-dimensional, it supplies only a single adjustable parameter, so the final term on the right could equally well be expressed $T^{12}\boldsymbol{\epsilon}_{12}$ or—more simply—$P\boldsymbol{\epsilon}$. So we write

$$\boldsymbol{G} = S\mathbf{I} + V^i\boldsymbol{\epsilon}_i + P\boldsymbol{\epsilon} \tag{42}$$

If

$$\boldsymbol{H} = s\mathbf{I} + v^i\boldsymbol{\epsilon}_i + p\boldsymbol{\epsilon}$$

describes an arbitrary second element in $\mathcal{C}\ell_2(g)$ then a little work supplies the product formula

$$\begin{aligned}\boldsymbol{GH} = (Ss + V^iv_i - gPp)\mathbf{I} + (Sv^j + sV^j - Pv_i\varepsilon^{ij} + pV_i\varepsilon^{ij})\boldsymbol{\epsilon}_j \\ + (Sp + V^1v^2 - V^2v^1 + Ps)\boldsymbol{\epsilon}\end{aligned} \tag{43}$$

where it is actually the expressions on the right sides of

$$\begin{aligned} V^iv_i &= V^1g_{11}v^1 + V^1g_{12}v^2 + V^2g_{21}v^1 + V^2g_{22}v^2 \\ g &= g_{11}g_{22} - g_{12}g_{21} \\ V_i &= g_{i1}V^1 + g_{i2}V^2 \\ v_i &= g_{i1}\,v^1 + g_{i2}\,v^2 \end{aligned}$$

which are presented to us in the course of the calculation, and where

$$\varepsilon^{ij} \equiv \operatorname{sgn}\tbinom{i\,j}{1\,2}$$

I strongly urge my reader to *do* the calculation that gives (43); it takes only a few minutes of careful work on a large sheet of paper, and is highly instructive.

[9] See "Aspects of Clifford algebras" (text of a seminar presented on 27 March 1967 to the Reed College Math Club—back in the days before there were such things Thursday Math Seminars) in COLLECTED SEMINARS (1963–1970).

If we define

$$\boldsymbol{G}^* = S\mathbf{I} - V^i\boldsymbol{\epsilon}_i - P\boldsymbol{\epsilon} \tag{44}$$

then it is an immediate implication of (43) that

$$\boldsymbol{G}^*\boldsymbol{G} = (S^2 - g_{ij}V^iV^j + gP^2)\mathbf{I}$$

Evidently $\boldsymbol{G}$ is invertible if and only if its "modulus"

$$|\boldsymbol{G}| \equiv S^2 - g_{ij}V^iV^j + gP^2 \neq 0 \tag{45}$$

In a matrix representation $\boldsymbol{G} \to \mathbb{G}$ the role of the modulus would be taken over by $\det \mathbb{G}$. We find it natural, therefore, to construct the

$$\text{"characteristic polynomial"}\ |\boldsymbol{G} - \lambda\mathbf{I}| = \lambda^2 - 2S\lambda + |\boldsymbol{G}|$$

and are not surprised to discover that $\boldsymbol{G}$ satisfies its own characteristic equation:

$$\boldsymbol{G}^2 - 2S\boldsymbol{G} + |\boldsymbol{G}| = \mathbf{0} \tag{46}$$

We are assured that $\mathcal{C}\ell_2(g)$ —since an associative algebra of order 4—admits of real 4×4 representation, but the preceding results suggest strongly that it admits actually of 2×2 representation.

The Clifford algebras $\mathcal{C}\ell_{2,0}$, $\mathcal{C}\ell_{1,1}$ and $\mathcal{C}\ell_{0,2}$ can be recovered as special cases of $\mathcal{C}\ell_2(g)$. To see how this works we look to the latter; i.e., to the quaternionic case

$$\|g_{ij}\| = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

After notational adjustments

$$\begin{array}{llll} S \to a^0 & s \to b^0 & & \\ V^1 \to a^1 & v^1 \to b^1 & \boldsymbol{\epsilon}_1 \to \boldsymbol{i} & \\ V^2 \to a^2 & v^2 \to b^2 & \boldsymbol{\epsilon}_2 \to \boldsymbol{j} & \\ P \to a^3 & p \to b^3 & \boldsymbol{\epsilon} \to \boldsymbol{k} & \end{array}$$

the general product formula (43) becomes

$$\begin{aligned} (a^0\mathbf{I} + a^1\boldsymbol{i} + a^2\boldsymbol{j} + a^3\boldsymbol{k})(b^0\mathbf{I} + b^1\boldsymbol{i} + b^2\boldsymbol{j} + b^3\boldsymbol{k}) & \\ = (a^0b^0 - a^1b^1 - a^2b^2 - a^3b^3)\mathbf{I} + (a^0b^1 + a^1b^0 + a^2b^3 - a^3b^2)\boldsymbol{i} & \\ + (a^0b^2 + a^2b^0 + a^3b^1 - a^1b^3)\boldsymbol{j} & \\ + (a^0b^3 + a^3b^0 + a^1b^2 - a^2b^1)\boldsymbol{k} & \end{aligned}$$

which is familiar as the *quaternionic product formula*—historic birthplace of the "dot product," the "cross product" (set $a^0 = b^0 = 0$) and of non-commutative algebra generally.

A notational adjustment

$$\mathsf{I} \to \boldsymbol{e}_0 \; ; \quad \boldsymbol{\epsilon}_1 \to \boldsymbol{e}_1 \; ; \quad \boldsymbol{\epsilon}_2 \to \boldsymbol{e}_2 \; ; \quad \boldsymbol{\epsilon} \to \boldsymbol{e}_3$$

places us in position to use (31) to construct a real 4×4 representation of $\mathcal{C}\ell_2(g)$. Working from (43) we have

$$\begin{aligned}
\boldsymbol{e}_0(s\boldsymbol{e}_0 + v^1\boldsymbol{e}_1 + v^2\boldsymbol{e}_2 + p\boldsymbol{e}_3) &= s\boldsymbol{e}_0 + v^1\boldsymbol{e}_1 + v^2\boldsymbol{e}_2 + p\boldsymbol{e}_3 \\
\boldsymbol{e}_1(s\boldsymbol{e}_0 + v^1\boldsymbol{e}_1 + v^2\boldsymbol{e}_2 + p\boldsymbol{e}_3) &= g_{1k}v^k\boldsymbol{e}_0 + (s - pg_{12})\boldsymbol{e}_1 + pg_{11}\boldsymbol{e}_2 + v^2\boldsymbol{e}_3 \\
\boldsymbol{e}_2(s\boldsymbol{e}_0 + v^1\boldsymbol{e}_1 + v^2\boldsymbol{e}_2 + p\boldsymbol{e}_3) &= g_{2k}v^k\boldsymbol{e}_0 - pg_{22}\boldsymbol{e}_1 + (s + pg_{21})\boldsymbol{e}_2 - v^1\boldsymbol{e}_3 \\
\boldsymbol{e}_3(s\boldsymbol{e}_0 + v^1\boldsymbol{e}_1 + v^2\boldsymbol{e}_2 + p\boldsymbol{e}_3) &= -gp\boldsymbol{e}_0 + v^i g_{i2}\boldsymbol{e}_1 - v^i g_{i1}\boldsymbol{e}_2 + s\boldsymbol{e}_3
\end{aligned}$$

giving

$$\mathsf{I} \;=\; \boldsymbol{e}_0 \to \;\mathbb{I} \;=\; \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{47.0}$$

$$\boldsymbol{\epsilon}_1 = \boldsymbol{e}_1 \to \mathbb{E}_1 = \begin{pmatrix} 0 & g_{11} & g_{12} & 0 \\ 1 & 0 & 0 & -g_{12} \\ 0 & 0 & 0 & g_{11} \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{47.1}$$

$$\boldsymbol{\epsilon}_2 = \boldsymbol{e}_2 \to \mathbb{E}_2 = \begin{pmatrix} 0 & g_{21} & g_{22} & 0 \\ 0 & 0 & 0 & -g_{22} \\ 1 & 0 & 0 & g_{21} \\ 0 & -1 & 0 & 0 \end{pmatrix} \tag{47.2}$$

$$\boldsymbol{\epsilon} \;=\; \boldsymbol{e}_3 \to \mathbb{E} \;=\; \begin{pmatrix} 0 & 0 & 0 & -g \\ 0 & g_{12} & g_{22} & 0 \\ 0 & -g_{11} & -g_{21} & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \tag{47.3}$$

Implausible though the design of these matrices may appear, *Mathematica* quickly confirms that

$$\mathbb{E}_i\,\mathbb{E}_j + \mathbb{E}_j\,\mathbb{E}_i = 2g_{ij}\,\mathbb{I} \quad \text{and} \quad \mathbb{E}_1\,\mathbb{E}_2 - g_{12}\,\mathbb{I} = \mathbb{E} \tag{48}$$

The representation of $\boldsymbol{G} = S\mathsf{I} + V^i\boldsymbol{\epsilon}_i + P\boldsymbol{\epsilon}$ becomes

$$\begin{aligned}
\mathbb{G} &= S\,\mathbb{I} + V^1\,\mathbb{E}_1 + V^2\,\mathbb{E}_2 + P\,\mathbb{E} \\
&= \begin{pmatrix} S & V_1 & V_2 & -Pg \\ V^1 & S+Pg_{12} & Pg_{22} & -V_2 \\ V^2 & -Pg_{11} & S-Pg_{21} & V_1 \\ P & -V^2 & V^1 & S \end{pmatrix}
\end{aligned} \tag{49}$$

and, appealing again to *Mathematica* for assistance, we find

$$\det\mathbb{G} = (S^2 - V^iV_i + gP^2)^2 = |\boldsymbol{G}|^2 \tag{50}$$

If, in (49), we set $S = P = 0$ we arrive back at our point of departure

$$(V^1\mathbb{E}_1 + V^2\mathbb{E}_2)^2 = \begin{pmatrix} 0 & V_1 & V_2 & 0 \\ V^1 & 0 & 0 & -V_2 \\ V^2 & 0 & 0 & V_1 \\ 0 & -V^2 & V^1 & 0 \end{pmatrix}^2$$
$$= (V_1V^1 + V_2V^2)\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

. . . which is comforting, if not surprising.

It should be borne in mind that the representation (47/49) is not unique: similarity transformations

$$\mathbb{E}_\mu \to \mathbb{E}_\mu \equiv \mathbb{T}^{-1}\mathbb{E}_\mu\mathbb{T} \tag{51}$$

yield alternative/equivalent representations.

A "real 4-component formulation of a toy Dirac theory with arbitrary metric" is our present objective, but before we can make further progress we must gain a sharper sense of certain general features shared by all "real-field Dirac theories," and of the conditions under which such theories admit of Lagangian formulation. We must, in particular, discover what becomes of the i which enters so conspicuously into the Dirac equation, but can have no place in a real-field theory. To get a handle on the points at issue we look back again to (5); i.e., to our toy theory in the Lorentzian case. If we

$$\text{let } \psi = \begin{pmatrix} \psi^1 \\ \psi^2 \end{pmatrix} = \begin{pmatrix} \alpha^1 + i\beta^1 \\ \alpha^2 + i\beta^2 \end{pmatrix} \text{ be expanded } \begin{pmatrix} \alpha^1 \\ \beta^1 \\ \alpha^2 \\ \beta^2 \end{pmatrix}$$

then (5) becomes (note the absence of i-factors!)

$$\left[\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\partial_0 + \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}\partial_1 + \varkappa\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}\right]\begin{pmatrix} \alpha^1 \\ \beta^1 \\ \alpha^2 \\ \beta^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

which will be abbreviated

$$\big[\boldsymbol{\Gamma}^0\partial_0 + \boldsymbol{\Gamma}^1\partial_1 + \varkappa\,\mathbb{J}\,\big]\Psi = 0 \tag{52.1}$$

We observe that

$$\boldsymbol{\Gamma}^\mu\boldsymbol{\Gamma}^\nu + \boldsymbol{\Gamma}^\nu\boldsymbol{\Gamma}^\mu = 2g^{\mu\nu}\mathbb{I} \quad \text{with} \quad \|g^{\mu\nu}\| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{52.2}$$

and—which are more to the point—that

$$\mathbb{J} \text{ commutes with } \boldsymbol{\Gamma}^0 \text{ and } \boldsymbol{\Gamma}^1\text{; moreover } \mathbb{J}^2 = -\mathbb{I} \tag{52.3}$$

In the complex Dirac theory of the textbooks one achieves (52.3) by setting $\mathbb{J} = i\,\mathbb{I}$, but that option is not in all cases forced...and in real theory not possible; $\mathbb{J}$ is a "square root of minus $\mathbb{I}$" in an enlarged, matrix-theoretic sense.

If (52.1) is to admit of Lagrangian formulation then there must exist a real non-singular matrix $\mathbb{S}$ such that

$$\mathbb{S}\boldsymbol{\Gamma}^0 \text{ and } \mathbb{S}\boldsymbol{\Gamma}^1 \text{ are antisymmetric, and } \mathbb{S}\mathbb{J} \text{ symmetric} \tag{52.4}$$

in which case we have[10]

$$\mathfrak{L} = -\tfrac{1}{2}\hbar c\big\{\Psi^{\intercal}\mathbb{S}\boldsymbol{\Gamma}^{\mu}\Psi_{,\mu} + \varkappa\Psi^{\intercal}\mathbb{S}\mathbb{J}\,\Psi\big\} \tag{52.5}$$

Exploratory tinkering (I know of no systematic method) shows that a $\mathbb{S}$ that does the job is

$$\mathbb{S} \equiv \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \tag{52.6}$$

for

$$\mathbb{S}\boldsymbol{\Gamma}^0 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \quad \text{and} \quad \mathbb{S}\boldsymbol{\Gamma}^1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

are antisymmetric, while

$$\mathbb{S}\mathbb{J} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

is symmetric.

Look now to the generality of the points at issue. The "Dirac factorization problem"

$$(\boldsymbol{\gamma}^{\mu}\partial_{\mu} - \varkappa\,\mathsf{J})(\boldsymbol{\gamma}^{\mu}\partial_{\mu} + \varkappa\,\mathsf{J}) = (\Box + \varkappa^2)\,\mathsf{I} \tag{53.1}$$

[10] Notice that the symmetric part of $\mathbb{S}\boldsymbol{\Gamma}^{\mu}$, if present, could be discarded as a gauge term. And that while in quantum theory with complex ψ it is permissible (and standard) to treat ψ and ψ^* as though they were formally independent, it would be senseless to assign "formal independence" to Ψ and $\Psi^{\intercal}$. Only when these points are understood does it become permissible (if not very useful) to write

$$\tilde{\Psi} \equiv \Psi^{\intercal}\mathbb{S}$$

requires, in addition to the familiar conditions $\boldsymbol{\gamma}^\mu\boldsymbol{\gamma}^\nu+\boldsymbol{\gamma}^\nu\boldsymbol{\gamma}^\mu = 2g^{\mu\nu}\mathsf{I}$, that

$$\mathsf{J} \text{ commutes with all } \boldsymbol{\gamma}\text{-matrices; moreover } \mathsf{J}^2 = -\mathsf{I} \tag{53.2}$$

From $[\mathsf{J},\boldsymbol{\gamma}^\mu] = \mathbf{0}$ it follows that J commutes with all elements of the Clifford algebra *generated* by the $\boldsymbol{\gamma}$-matrices; i.e., that

$$\mathsf{J} \in \text{ the "center" of the algebra} \tag{53.3}$$

In standard complex-field Dirac theory the center contains only multiples of I, and the introduction of i-factors is forced. But in real-field Dirac theory the center may/must contain additional elements. The Dirac equation

$$(\boldsymbol{\gamma}^\mu\partial_\mu + \varkappa\,\mathsf{J})\Psi = 0 \tag{53.4}$$

admits of Lagrangian formulation if and only if there exists an invertible $\boldsymbol{S}$ such that

$$\boldsymbol{S}\boldsymbol{\gamma}^\mu \text{ is antisymmetric (all } \mu) \text{ and } \boldsymbol{S}\mathsf{J} \text{ is symmetric} \tag{53.5}$$

in which case one has

$$\mathcal{L} = -\tfrac{1}{2}\hbar c\big\{\Psi^\intercal\boldsymbol{S}\boldsymbol{\gamma}^\mu\Psi_{,\mu} + \varkappa\Psi^\intercal\boldsymbol{S}\mathsf{J}\Psi\big\} \tag{53.6}$$

But what we presently lack are constructive means to *exhibit* matrices J and $\boldsymbol{S}$ with the requisite properties, or even guarantee of their existence. That this is a major handicap will soon be evident:

In the Lorentzian case the regular representation formulae (47.1) and (47.2) supply[11]

$$\mathit{\Gamma}^1 = g^{1k}\mathbb{E}_k = +\mathbb{E}_1 = \begin{pmatrix} 0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix}$$

$$\mathit{\Gamma}^2 = g^{2k}\mathbb{E}_k = -\mathbb{E}_2 = \begin{pmatrix} 0 & 0 & 1 & 0\\ 0 & 0 & 0 & -1\\ -1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{pmatrix}$$

which when combined with $\mathit{\Gamma} \equiv \mathit{\Gamma}^{12} \equiv \mathit{\Gamma}^1\mathit{\Gamma}^2 - g^{12}\,\mathbb{I} = \mathit{\Gamma}^1\mathit{\Gamma}^2$ entail

$$\begin{aligned}\mathbb{G} &= S\mathbb{I} + V_1\mathit{\Gamma}^1 + V_2\mathit{\Gamma}^2 + P\mathit{\Gamma}\\ &= \begin{pmatrix} S & V_1 & V_2 & -P\\ V_1 & S & P & -V_2\\ -V_2 & P & S & V_1\\ -P & V_2 & V_1 & S\end{pmatrix}\end{aligned}$$

[11] The following remarks are intended to be read on-screen, and will be rendered confusing by loss of the colored typography which I have used to avoid confusing hats, primes, etc.

Notice that $\mathbb{\Gamma}^0$ and $\mathbb{\Gamma}^1$ differ from the $\mathbb{\Gamma}^0$ and $\mathbb{\Gamma}^1$ introduced at (52.1). I have not been able to establish their equivalence in the sense (51), and suspect that they are *not* equivalent. Nor have I been able to discover either a $\mathbb{J}$ which commutes with $\mathbb{\Gamma}^0$ and $\mathbb{\Gamma}^1$ or a $\mathbb{S}$ which renders $\mathbb{S}\mathbb{\Gamma}^0$ and $\mathbb{S}\mathbb{\Gamma}^1$ antisymmetric (but neither have I been able to show that such things are impossible). Multiple failure in this relatively simple case leads me to think that it will be difficult/ impossible to erect a "Lagrangian formulation of real 4-component toy Dirac theory with general metric" on the platform provided by (49).

But why are we interested in real 4-component theory? Only because the regular representation of $\mathcal{C}\ell_2(g)$ has supplied real 4×4 matrices $\mathbb{\Gamma}^0$ and $\mathbb{\Gamma}^1$. In the thought that "general metric theory" should be constructed on some alternative pattern, we observe that if we introduce

$$\mathbb{H} \equiv \begin{pmatrix} 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \quad : \quad \text{hermitian} \tag{54.1}$$

then (49) supplies

$$\mathbb{HG} = S\mathbb{H} + \underbrace{i\begin{pmatrix} P & -V^2 & +V^1 & 0 \\ -V^2 & Pg_{11} & Pg_{12} & -V_1 \\ +V^1 & Pg_{21} & Pg_{22} & -V_2 \\ 0 & -V_1 & -V_2 & Pg \end{pmatrix}}_{\text{antihermitian!}} \tag{54.2}$$

The pretty implication is that if we take (47.1) and (47.2) to be definitions of $\mathbb{\Gamma}_0 \equiv \mathbb{E}_0$ and $\mathbb{\Gamma}_1 \equiv \mathbb{E}_1$ then

$$(\mathbb{\Gamma}_\mu\partial^\mu - i\varkappa\mathbb{I})(\mathbb{\Gamma}_\nu\partial^\nu + i\varkappa\mathbb{I}) = (\Box + \varkappa^2)\mathbb{I}$$
$$\Box \equiv g_{\mu\nu}\partial^\mu\partial^\nu \text{ with general metric}$$

which motivates us to write

$$(\mathbb{\Gamma}_\mu\partial^\mu + i\varkappa\mathbb{I})\psi = 0 \quad \text{with} \quad \psi = \begin{pmatrix} \psi^1 \\ \psi^2 \\ \psi^3 \\ \psi^4 \end{pmatrix} \tag{55}$$

and to observe that this "toy Dirac equation with general metric" can be obtained from a Lagrangian of (compare (9.0)) this classic design:

$$\mathfrak{L} = -\hbar c\Big[i\tfrac{1}{2}\{\psi^\dagger_{,\mu}\mathbb{H}\mathbb{\Gamma}^\mu\psi - \psi^\dagger\mathbb{H}\mathbb{\Gamma}^\mu\psi_{,\mu}\} + \varkappa\psi^\dagger\mathbb{H}\psi\Big] \tag{56}$$

Were one to separate the real from the imaginary components of ψ one would be led in the toy theory to an 8-component wave function Ψ (and in the physical case studied by Dirac to a 32-component wave function!).

Anticipated continuation. In the 4-component theory (55) the $\boldsymbol{\Gamma}$-matrices are 4×4 and real. When I have an opportunity to resume this discussion I will attempt to construct a 2×2 complex realization of $\mathcal{C}\ell_2(g)$. While it was established by Dirac that $\mathcal{C}\ell_{1,3}$ admits of 4×4 complex realization, it is not clear that $\mathcal{C}\ell_4(g)$ does; if so, then one might possibly expect to have a 3×3 complex realization $\mathcal{C}\ell_3(g)$ of, even though the regular realization is 8×8. It is my experience, however, that—in this area especially—numerology is an unreliable guide.

I propose to explore also the following topics:

• Foldy-Wouthuysen representation in 2-dimensional theory (take maybe Schweber's §4f as my point of departure).

• Massless Dirac fields in 2-dimensional theory.

• Abelian/non-Abelian gauge field theories supported by the 2-dimensional Dirac theory.

• Dirac theory on curved 2-dimensional manifolds $g_{\mu\nu}(x)$. The "Vierbein formalism" becomes a "Zweibein formalism"? Use general covariance to achieve symmetrization of the stress energy tensor?

• At (54.2) we encounter (set $S=V^1=V^2=0$ and $P=1$) a matrix of the form

$$\begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & g_{11} & g_{12} & 0\\ 0 & g_{21} & g_{22} & 0\\ 0 & 0 & 0 & g \end{pmatrix}$$

encountered also in exterior algebra (See ELECTRODYNAMICS (1972), p. 156). It would be instructive to develop the connection (which is well known to people like Lounesto).

THE ANISOTROPIC STRING

Nicholas Wheeler, Reed College Physics Department
December 2004

Introduction. To describe the motion of waves on a stretched string we write

$$\left\{\partial_x^2 - \tfrac{1}{uu}\partial_t^2\right\}\varphi(x,t) = 0 \tag{1}$$

which (since the predominant waves on a string are transverse, and their description requires therefore that we monitor *two* fields) relates more properly to compressional waves on a stretched spring. Factorization of the wave operator

$$\partial_x^2 - \tfrac{1}{uu}\partial_t^2 = \left(\partial_x + \tfrac{1}{u}\partial_t\right)\left(\partial_x - \tfrac{1}{u}\partial_t\right)$$

leads directly to the conclusion that φ can be developed

$$\varphi(x,t) = f(x-ut) + g(x+ut)$$

where the right-running wave $f(x-ut)$ is killed by $\left(\partial_x + \frac{1}{u}\partial_t\right)$, the left-running wave $g(x+ut)$ is killed by $\left(\partial_x - \frac{1}{u}\partial_t\right)$.

Which brings us to the simple point of this note: Suppose it were the case that waves run right and left with *distinct speeds* u and v. We would then be led to write

$$\varphi(x,t) = f(x-ut) + g(x+vt)$$

which is a solution of

$$\left(\partial_x + \tfrac{1}{u}\partial_t\right)\left(\partial_x - \tfrac{1}{v}\partial_t\right)\varphi = \left\{\partial_x^2 + \left(\tfrac{1}{u} - \tfrac{1}{v}\right)\partial_x\partial_t - \tfrac{1}{uv}\partial_t^2\right\}\varphi = 0 \tag{2}$$

We observe that (2) gives back (1) in the case $u = v$. My intention is to examine some of the formal/physical properties of the anisotropic system (2).

1. Lagrangian formalism. Introduce

$$\mathcal{L}(\varphi,\partial\varphi) = \tfrac{1}{2uv}\varphi_t^2 - \tfrac{1}{2}\left(\tfrac{1}{u} - \tfrac{1}{v}\right)\varphi_x\varphi_t - \tfrac{1}{2}\varphi_x^2 \tag{3}$$

and from

$$\left\{\partial_t\frac{\partial}{\partial\varphi_t} + \partial_x\frac{\partial}{\partial\varphi_x} - \frac{\partial}{\partial\varphi}\right\}\mathcal{L} = 0$$

obtain

$$\tfrac{1}{uv}\varphi_{tt} - \left(\tfrac{1}{u} - \tfrac{1}{v}\right)\varphi_{tx} - \varphi_{xx} = 0$$

—which is (2). The Lagrangian (3) gives rise to a stress-energy tensor the components of which can be described[1]

$$\left.\begin{aligned}
S^t{}_t &= \frac{\partial\mathcal{L}}{\partial\varphi_t}\varphi_t - \mathcal{L}\\
&= \tfrac{1}{2uv}\varphi_t^2 + \tfrac{1}{2}\varphi_x^2\\
&\equiv \text{energy density } \mathcal{E}\\[2ex]
S^x{}_t &= \frac{\partial\mathcal{L}}{\partial\varphi_x}\varphi_t\\
&= -\tfrac{1}{2}\big(\tfrac{1}{u}-\tfrac{1}{v}\big)\varphi_t^2 - \varphi_x\varphi_t\\
&\equiv \text{energy flux } \mathcal{F}\\[2ex]
S^t{}_x &= \frac{\partial\mathcal{L}}{\partial\varphi_t}\varphi_x\\
&= \tfrac{1}{uv}\varphi_t\varphi_x - \tfrac{1}{2}\big(\tfrac{1}{u}-\tfrac{1}{v}\big)\varphi_x^2\\
&\equiv \text{momentum density } \mathcal{P}\\[2ex]
S^x{}_x &= \frac{\partial\mathcal{L}}{\partial\varphi_x}\varphi_x - \mathcal{L}\\
&= -\tfrac{1}{2uv}\varphi_t^2 - \tfrac{1}{2}\varphi_x^2\\
&\equiv \text{momentum flux } \mathcal{G}
\end{aligned}\right\} \tag{4}$$

By calculation we verify that

$$\begin{aligned}
\partial_t S^t{}_t + \partial_x S^x{}_t &= \varphi_t\big[\tfrac{1}{uv}\varphi_{tt} - \big(\tfrac{1}{u}-\tfrac{1}{v}\big)\varphi_{tx} - \varphi_{xx}\big] = 0\\
\partial_t S^t{}_x + \partial_x S^x{}_x &= \varphi_x\big[\tfrac{1}{uv}\varphi_{tt} - \big(\tfrac{1}{u}-\tfrac{1}{v}\big)\varphi_{tx} - \varphi_{xx}\big] = 0
\end{aligned}$$

which are statements of local energy/momentum conservation.

2. The equivalent lattice. Consider the familiar "one-dimensional crystal" that has been assembled from identical particles—each of mass m, each coupled to its nearest neighbors by identical springs of elasticity k. To describe the instantaneous position of the n^{th} particle we write

$$\begin{aligned}
x_n(t) &= na + \varphi_n(t)\\
&= \text{equilibrium position} + \text{displacement}
\end{aligned}$$

where a is the "lattice constant." To describe the motion of the n^{th} particle (unless it is an end-particle) we write

$$\begin{aligned}
m\ddot{\varphi}_n &= k(\varphi_{n+1}-\varphi_n) - k(\varphi_n - \varphi_{n-1})\\
&= -k(-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1})
\end{aligned} \tag{4}$$

[1] See CLASSICAL FIELD THEORY (1999), Chapter 1, page 63.

This coupled system of equations can be notated

$$\mathbb{M}\,\ddot{\boldsymbol{\varphi}} + \mathbb{K}\boldsymbol{\varphi} = \mathbf{0} \tag{6}$$

with $\mathbb{M} \equiv m\mathbb{I}$ and

$$\mathbb{K} \equiv k \begin{pmatrix} * & * & & & & & * \\ \alpha & \beta & \alpha & & & & \\ & \alpha & \beta & \alpha & & & \\ & & \alpha & \beta & \alpha & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \alpha & \beta & \alpha \\ * & & & & & * & * \end{pmatrix}$$

where $\alpha = -1$, $\beta = +2$, the values assigned to the $*$s depends upon how we elect to manage the end-particles (coupled to fixed walls? coupled to each other?), and all other elements are zero. We observe that $\mathbb{M}$ and $\mathbb{K}$ are both symmetric, and that (6) can be obtained from a Lagrangian of the design

$$L_0 = \tfrac{1}{2}\dot{\boldsymbol{\varphi}}\cdot\mathbb{M}\dot{\boldsymbol{\varphi}} + \tfrac{1}{2}\boldsymbol{\varphi}\cdot\mathbb{K}\boldsymbol{\varphi}$$

Standard stuff. But we take now a non-standard step: we introduce into the Lagrangian a "gyroscopic term," writing

$$L = \tfrac{1}{2}\dot{\boldsymbol{\varphi}}\cdot\mathbb{M}\dot{\boldsymbol{\varphi}} + \tfrac{1}{2}\dot{\boldsymbol{\varphi}}\cdot\mathbb{G}\boldsymbol{\varphi} + \tfrac{1}{2}\boldsymbol{\varphi}\cdot\mathbb{K}\boldsymbol{\varphi} \tag{7}$$

where $\mathbb{G}$ is an *anti*symmetric matrix of the design

$$\mathbb{G} \equiv g \begin{pmatrix} 0 & -1 & & & \\ +1 & 0 & -1 & & \\ & +1 & 0 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & & \end{pmatrix}$$

The equations of motion (6)—thus adjusted—become

$$\mathbb{M}\,\ddot{\boldsymbol{\varphi}} + \mathbb{G}\dot{\boldsymbol{\varphi}} + \mathbb{K}\boldsymbol{\varphi} = \mathbf{0}$$

which in fine detail (compare (4)) read

$$m\ddot{\varphi}_n = g(\dot{\varphi}_{n+1} - \dot{\varphi}_{n-1}) - k(-\varphi_{n-1} + 2\varphi_n - \varphi_{n+1}) \tag{8}$$

We want now to "refine the lattice"—to make the particles progressively more numerous, individually (though not collectively) less massive, more closely spaced and to regulate the values of g and k in such a way as to obtain a meaningful continuous (or "field-theoretic") limit of the discrete system

presently in hand. The program[2] proceeds from a notational adjustment: in place of $\varphi_n(t)$ we write $\varphi(na,t)$, which in the continuous limit becomes $\varphi(x,t)$ and will be abbreviated $\varphi(x)$ when t is not a participant but merely a spectator. In that notation (8) becomes

$$\varphi_{tt}(x) = \frac{2g(a)\,a}{m(a)} \cdot \frac{\varphi_t(x+a) - \varphi_t(x+a)}{2a} + \frac{k(a)\,a^2}{m(a)} \cdot \frac{\frac{\varphi(x+a)-\varphi(x)}{a} - \frac{\varphi(x)-\varphi(x-a)}{a}}{a}$$

To recover (2)—*i.e.*, to recover

$$\varphi_{tt} - (v-u)\varphi_{tx} - uv\,\varphi_{xx} = 0$$

—in the limit $a \downarrow 0$ we have only to stipulate that

$$\lim_{a\downarrow 0} \frac{2g(a)\,a}{m(a)} = v - u \qquad \text{and} \qquad \lim_{a\downarrow 0} \frac{k(a)\,a^2}{m(a)} = uv \tag{9}$$

Since

$$\frac{m(a)}{a} \text{ models linear mass density } \mu$$

we are at (9) requiring in effect that

$$g(a) = \tfrac{1}{2}\mu \cdot (v-u)$$

remains constant as a diminishes, while the springs get stiffer as they get shorter, and become infinitely stiff in the limit[3]

$$a\,k(a) = \mu uv$$

The Lagrangian of the discrete system (which for us has become an a-parameterized *sequence* of discrete systems)—which at (7) is presented as a sum of terms—is readily seen to go over in the continuous limit into an integral:

$$L = \int \mathcal{L}\,dx$$
$$\mathcal{L} = \tfrac{1}{2}\mu\varphi_t^2 - \tfrac{1}{2}\mu(v-u)\varphi_t\varphi_x - \tfrac{1}{2}\mu uv\,\varphi_x^2$$

Division by μuv gives back the Lagrange density encountered at (3).

---

[2] See pages 5–8 in the class notes just cited, or §13.1 in H. Goldstein *et al*, *Classical Mechanics* (3rd edition 2002).

[3] This development is not at all "strange:" it is basic that springs get softer when connected in series, stiffer when chopped into fragments.

In his Reed College thesis[4] Mark Galassi demonstrated that lattices of the design

$$ABCABCABCABC\cdots$$

do *not* support anisotropic wave physics, even though they present distinct faces to right-moving and left-moving tourists. What we have learned is that anisotropy is a "gyroscopic" effect,[5] that it arises from a velocity-dependent "gyroscopic coupling" of particles or string-elements to their (near) neighbors.

It was established on page 8 of the material just cited that gyroscopic terms have no effect upon the energetics of oscillatory systems. At (4) we found that such terms have no effect either upon the energy density of anisotropic (or as we now understand them to be: gyroscopic) strings. But they do show up in the formulæ that describe energy *flux* and *momentum* density.

3. Transformational aspects of anisotropy. If we, in the rest frame of the string, see waves to run $\rightarrow$ with speed u, $\leftarrow$ with speed v, then we expect an observer who is himself running $\rightarrow$ with speed c to see

- waves to run $\rightarrow$ with speed $u-c$
- waves to run $\leftarrow$ with speed $v+c$

and to see apparent isotropy in the case $c=\frac{1}{2}(u-v)$, the $\leftrightarrow$ wave speed in that case being

$$w=\tfrac{1}{2}(u+v)$$

Our moving observer has *transformationally eliminated the anisotropy*, which is consonant with the upshot of §3 in the material just cited, where it is established that "the gyroscopic term can always be rotated away."

It is easy to show[6] that under Galilean transformation

$$\boldsymbol{t}=t$$
$$\boldsymbol{x}=x+ct$$

the familiar isotropic wave equation

$$\Big\{\big(\tfrac{\partial}{\partial \boldsymbol{x}}\big)^2-\tfrac{1}{w^2}\big(\tfrac{\partial}{\partial \boldsymbol{t}}\big)^2\Big\}\boldsymbol{\varphi}(\boldsymbol{x},\boldsymbol{t})=0$$

goes over into

$$\Big\{\big(\tfrac{\partial}{\partial x}\big)^2+\tfrac{2c}{w^2-c^2}\tfrac{\partial}{\partial x}\tfrac{\partial}{\partial t}-\tfrac{1}{w^2-c^2}\big(\tfrac{\partial}{\partial t}\big)^2\Big\}\varphi(x,t)=0$$

But $w+c=u$, $w-c=v$ and $2c=u-v$ so we recover the equation

$$\varphi_{xx}-\big(\tfrac{1}{u}-\tfrac{1}{v}\big)\varphi_{tx}-\tfrac{1}{uv}\varphi_{tt}=0$$

that appears at the bottom of page 1. From this point of view the Lorentz transformations can be said to have been invented "to kill the gyroscopic term," and thus to turn isotropy into a frame-independent (observer-independent) concept.

---

[4] "Lagrangian field theory of anisotropic systems" (1986).

[5] See §§2&3 in Chapter 3 of ADVANCED CLASSICAL MECHANICS (2004).

[6] See "Electrodynamics in 2-dimensional spacetime" (1997), page 20.

4. Running planewaves. At the end of the preceding section a fresh trail led us back again to our starting point. So does this short trail:

Require of the function $e^{i(kx-\omega t)}$ that it be a solution of (2), written

$$uv\varphi_{xx} + (v-u)\varphi_{xt} - \varphi_{tt} = 0 \tag{10}$$

Immediately $-uvk^2 + (v-u)k\omega + \omega^2 = 0$ which gives

$$\omega = \tfrac{1}{2}[(u-v) \pm (u+v)]k = \begin{cases} +ku, \text{ else} \\ -kv \end{cases}$$

So we have right/left-running planewaves

$$e^{ik(x-ut)} \quad \text{and} \quad e^{ik(x+vt)}$$

All waves within each population run with the same phase velocity (u else $-v$), so superpositions

$$f(x-ut) = \int F(k)e^{ik(x-ut)}\,dk$$
$$g(x+vt) = \int G(k)e^{ik(x+vt)}\,dk$$

are non-dispersive.

A simple model of a dispersive anisotropic string results if into (10) we introduce a "Klein-Gordon term," writing

$$uv\varphi_{xx} + (v-u)\varphi_{xt} - \varphi_{tt} - uv\kappa^2\varphi = 0 \tag{10}$$

We then find

$$\omega = \frac{(u-v)k \pm \sqrt{(u+v)^2k^2 + 4uv\kappa^2}}{2} \tag{11}$$

In the case $u = v = c$ the field equation (10) assumes standard Klein-Gordon form

$$\varphi_{xx} - \tfrac{1}{c^2}\varphi_{tt} - \kappa^2\varphi = 0$$

and the dispersion equation (11) becomes

$$\omega/c = \pm\sqrt{k^2 + \kappa^2}$$

We conclude that anisotropy and dispersion are in any event not mutually exclusive. Indeed, we may expect their simultaneous presence to the rule rather than the exception, for—on the general grounds that very little in physics is *exactly* so—we cannot expect material media or even the vacuum to be exactly isotropic, however exquisite may be the approximation, nor can we expect even the photon to be exactly massless.