GEOMETRICAL MECHANICS ‡

Remarks Commemorative of Heinrich Hertz

Nicholas Wheeler, Reed College Physics Department
February 1994

Introduction. My remarks today derive from a conversation I had in October 1993 with a student in quest of a classical mechanics project—a conversation in which evidently I found my own speculative words to be of more interest that did my visitor. He, in any event, elected to explore some other problem, and I found myself in the position familiar to the Little Red Hen, doing a "project" for a course in which I was not even enrolled! I shall begin with an account of that conversation. How I got from that event to a renewed/deepened interest in the life and work of H. Hertz will emerge in due course.

1. "Transit time" in 1-dimensional mechanics. To describe (relative to an inertial frame) the 1-dimensional motion of a mass point m we were taught by Newton to write

$$m\ddot{x} = F(x)$$

If $F(x)$ is "conservative" $\qquad F(x) = -\frac{d}{dx}U(x)$

(which in the 1-dimensional case is automatic) then, by a familiar line of argument,

$$E \equiv \tfrac{1}{2}m\dot{x}^2 + U(x) \quad \text{is conserved:} \quad \dot{E} = 0$$

Therefore the speed of the particle when at x can be described

$$v(x) = \sqrt{\tfrac{2}{m}\big[E - U(x)\big]} \tag{1}$$

and is determined (see the figure) by the "depth $E - U(x)$ of the potential lake." Several useful conclusions are immediate. The motion of m is bounded $a \leq x \leq b$ by "turning points" a and b where the potential lake has vanishing depth, and excluded from "forbidden regions" where $E - U(x) < 0$; i.e., where the potential has risen above the "lake level" E. And the dynamical time of flight, or "transit time" $x_0 \longrightarrow x$ can be described

$$t(x; x_0, E) = \int_{x_0}^{x} \frac{1}{\sqrt{\tfrac{2}{m}\big[E - U(y)\big]}}\, dy \tag{2}$$

‡ Notes for a Reed College Physics Seminar presented 23 February 1994.

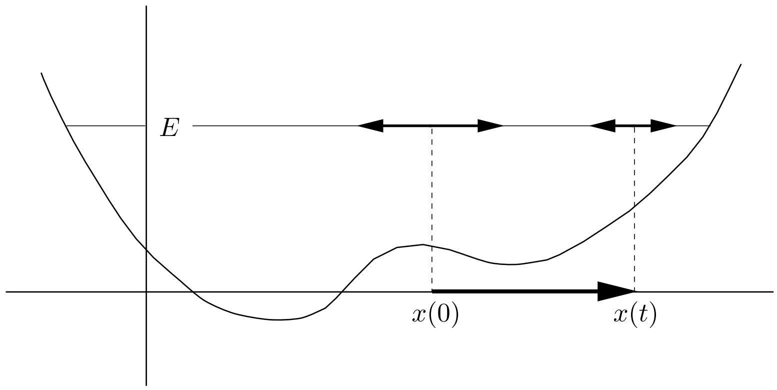

FIGURE 1: *"Potential lake" in 1-dimensional mechanics.*

By functional inversion—when it can be accomplished!—one obtains

$$x = x(t; x_0, E)$$

which provides an explicit description of the E-conserving motion of the particle.

Greater interest attaches, however, to (because they are less familiar) some of the *non*-standard applications/generalizations of (2). For example: functions of the type

$$T(x,p) = \int^x \frac{1}{\sqrt{\frac{2}{m}\big[H(x,p) - U(y)\big]}}\,dy$$

comprise seldom-remarked "natural companions" of the Hamiltonian

$$H(x,p) = \tfrac{1}{2m}p^2 + U(x)$$

Indeed, $H(x,p)$ and $T(x,p)$ are "conjugate observables" in the sense that

$$[T,H] \equiv \tfrac{\partial T}{\partial x}\tfrac{\partial H}{\partial p} - \tfrac{\partial H}{\partial x}\tfrac{\partial T}{\partial p} = 1$$

It is by quantization of the Hamiltonian $H(x,p) \longrightarrow \mathsf{H}$ that one prepares to write the Schrödinger equation $\mathsf{H}\psi = i\hbar\dot{\psi}$. The same procedure $T(x,p) \longrightarrow \mathsf{T}$ yields a highly non-standard object: a "time operator," which is conjugate to H in the sense standard to quantum mechanics

$$[\mathsf{T},\mathsf{H}] \equiv \mathsf{TH} - \mathsf{HT} = i\hbar\,\mathsf{I}$$

Upon this remark hangs a tale which I may tell on some other occasion. It is, however, by dimensional generalization of (2) that I was led to the subject matter of today's talk.

2. "Transit time" in N-dimensional mechanics. To describe (relative to an inertial frame) the N-dimensional motion of m we write

$$m\ddot{\boldsymbol{x}} = \boldsymbol{F}(\boldsymbol{x})$$

Now

$$\boldsymbol{F}(\boldsymbol{x}) = -\boldsymbol{\nabla} U(\boldsymbol{x})$$

is no longer automatic; it is mathematically rare but physically commonplace. It is, in all events, critical to energy conservation

$$\dot{E} = 0 \quad \text{with} \quad E \equiv \tfrac{1}{2}mv^2 + U(\boldsymbol{x})$$

and will be assumed. Speed v is, as before, determined by the local depth of the potential lake

$$v(\boldsymbol{x}) = \sqrt{\tfrac{2}{m}\big[E - U(\boldsymbol{x})\big]} \tag{3}$$

but now the potential lake is (see the second figure) a much more "lake-like"

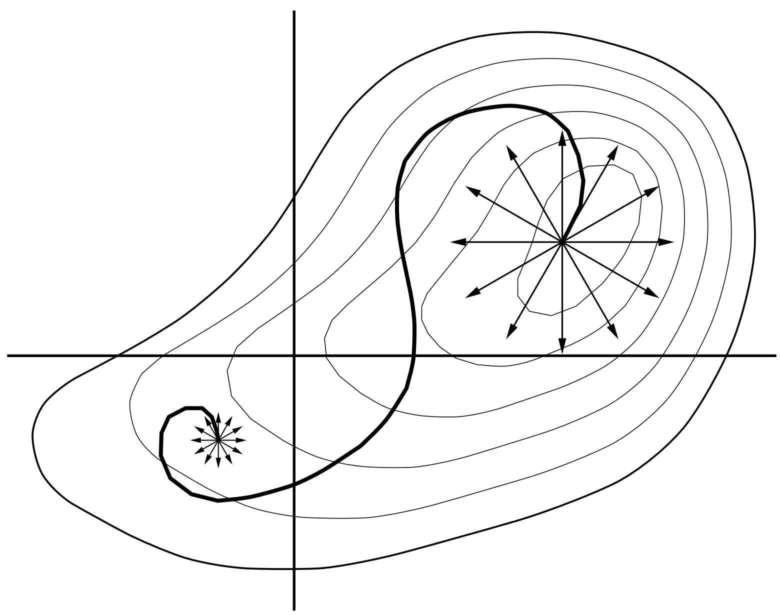

FIGURE 2: *"Topographic map" of a 2-dimensional potential lake. The light curves are equipotentials. The curve which links the points* $\boldsymbol{x}(0) \longrightarrow \boldsymbol{x}(s)$ *has Euclidean arc length* s.

place. It is, in particular, a place where *speed—since it conveys no directional information—is insufficient to determine velocity.*

To every "path" inscribed "on the surface of the E-lake" (i..e., within the region $\mathcal{L}$ bounded by the equipotential $U(\boldsymbol{x}) = E$) we can associate a "transit time" T[path]. To notate this obvious fact it is convenient to adopt arc-length parameterization $ds^2 = dx^2 + dy^2$, writing

$$T[\boldsymbol{x}(s)] = \int_0^s \frac{1}{v(\boldsymbol{x}(s'))}\,ds' \tag{4}$$

Consider now the population $\mathcal{P}$ of paths (of various lengths) $\boldsymbol{a} \longrightarrow \boldsymbol{b}$, as illustrated in the third figure. We have particular interest in the "dynamical" elements of such populations $\mathcal{P}$, i.e., in the paths which the E-conserving moton of m would trace out in time. More specifically, we have interest in the answer

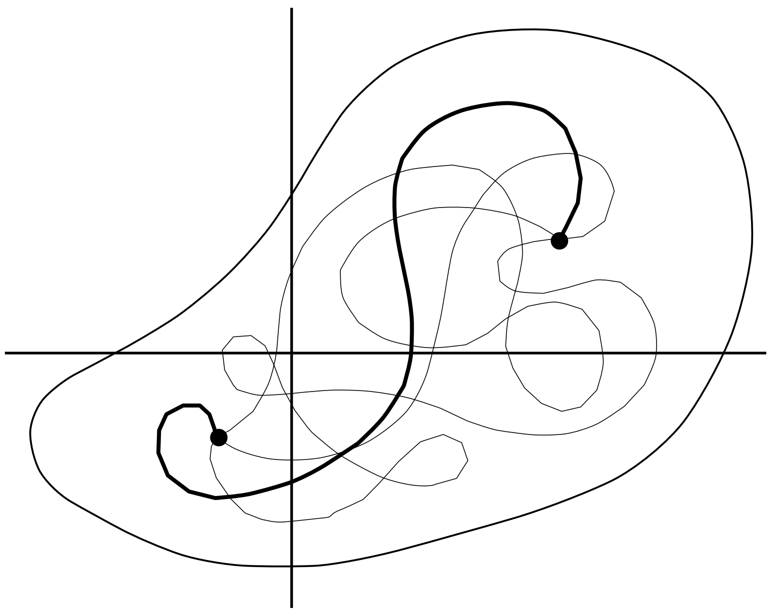

FIGURE 3: *Population $\mathcal{P}$ of curves inscribed on the surface of an E-lake. The distinguished curve was traced by a mass point m; it is "dynamically necessary," the others "dynamically impossible."*

to this question: *Is there a property of T[path] that serves to distinguish dynamical paths from paths-in-general?* The question—which is the question posed to the student mentioned in my Introduction—springs naturally to the imagination of anyone passingly familiar with the variational principles of mechanics, particularly because it is so evocative of . . .

3. Fermat's Principle of Least Time. A satisfactory geometrical optics—a theory of *light rays in isotropically inhomogeneous media* can, whether one proceeds from Maxwellian electrodynamics or in the more phenomenological language of Pierre Fermat, be constructed as follows:

1) To each point $\boldsymbol{x}$ in the medium assign a "speed function"

$$v(\boldsymbol{x}) = \frac{c}{n(\boldsymbol{x})}$$

where $n(\boldsymbol{x})$ is the local "index of refraction";

2) To each hypothetical path $\boldsymbol{x}(s)$ associate a number-valued path functional

$$\begin{aligned} T[\text{path}] &= \text{“transit time”} \\ &= \int_0^s \frac{1}{v(\boldsymbol{x}(s'))}\,ds' \\ &= \frac{1}{c}\cdot\underbrace{\int_0^s n(\boldsymbol{x}(s'))\,ds'}_{\equiv\ \text{“optical path length”}} \end{aligned}$$

3) Associate optical "rays" with the paths which extremize (or as the informal phrase goes, which "minimize") optical path length.

Turning now to the analytical implementation of Fermat's Principle, it proves convenient (to avoid a certain technical complication, as discussed below) to give up the specialness of s-parameterization in favor of unspecialized/ arbitrary λ-parameterization, writing $\boldsymbol{x}(\lambda)$ to describe a path. Then, using $ds = \sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}}\,d\lambda$ with $\mathring{\boldsymbol{x}} \equiv \frac{d}{d\lambda}\boldsymbol{x}$, we have

$$T[\boldsymbol{x}(\lambda)] = \tfrac{1}{c}\int n(\boldsymbol{x})\sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}}\,d\lambda$$

Then methods standard to the calculus of variations proceed from

$$\text{FERMAT'S PRINCIPLE:}\quad \delta T[\boldsymbol{x}(\lambda)] = 0$$

to the statement

$$\left\{\frac{d}{d\lambda}\frac{\partial}{\partial\mathring{\boldsymbol{x}}} - \frac{\partial}{\partial\boldsymbol{x}}\right\} n(\boldsymbol{x})\sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}} = \boldsymbol{0}$$

Thus are we led to the so-called "ray equations"

$$\frac{d}{d\lambda}\left[n\,\frac{\mathring{\boldsymbol{x}}}{\sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}}}\right] - \sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}}\,\boldsymbol{\nabla}n = \boldsymbol{0} \tag{5}$$

Several remarks are now in order:

If at this point we were to revert $s \longleftarrow \lambda$ to arc-length parameterization then $ds = d\lambda$ entails $\sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}} = 1$ and from (5) we obtain

$$\frac{d}{ds}\left[n\,\frac{d\boldsymbol{x}}{ds}\right] - \boldsymbol{\nabla}n = \boldsymbol{0} \tag{6}$$

which in homogeneous media (where $\boldsymbol{\nabla}n = \boldsymbol{0}$) reduces to $d^2\boldsymbol{x}/ds^2 = \boldsymbol{0}$: rays become straight (in the Euclidean sense).

Curiously, equation (6) is not itself derivable (except by trickery) from a "Lagrangian." To retain access to the Lagrangian method after adoption of s-parameterization one must treat $\sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}} = 1$ (here $\mathring{\boldsymbol{x}} \equiv \frac{d}{ds}\boldsymbol{x}$) as a "constraint," writing

$$\delta \int \left\{ n(\boldsymbol{x}) + \tfrac{1}{2}\lambda[\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}} - 1]\right\} = 0$$

where $\lambda(s)$ has now the nature of a "Lagrange multiplier" which has joined $\boldsymbol{x}(s)$ on the list of unknown functions of s which it is our business to describe.

Multiply $n(\boldsymbol{x})$ into (6) and obtain $\left(n\frac{d}{ds}\right)^2\boldsymbol{x} = \frac{1}{2}\boldsymbol{\nabla}n^2$. Write $n\frac{d}{ds} = \frac{d}{du}$; i.e., give up s-parameterization in favor of u-parameterization, with

$$u(s) = \int^s \frac{1}{n(\boldsymbol{x}(s'))}\, ds'$$

Then (6) reads

$$\left(\frac{d}{du}\right)^2 \boldsymbol{x} = \frac{1}{2}\boldsymbol{\nabla}n^2$$

which looks very "Newtonian," and can be considered to arise from the following "Lagrangian":

$$L = \frac{1}{2}\frac{d\boldsymbol{x}}{du}\cdot\frac{d\boldsymbol{x}}{du} + \frac{1}{2}n^2$$

Though Fermat spoke casually of "least time," what he actually gave us is a static *theory of curves*, in which "rays" are distinguished from other curves by their *least optical length*. It is, I think, well to remind ourselves that Fermat wrote in 1657, almost twenty years before Olaf Römer—in 1676, eleven years after Fermat's death—first demonstrated the speed of light to be $c < \infty$. We find it so utterly natural to think of the index of refraction as having to do with the ratio of two speeds

$$\text{index of refraction} = \frac{\text{speed of light in vacuum}}{\text{speed of light in medium}}$$

that we are astonished by the realization that neither Snell, nor Descarte, nor Fermat were in position to entertain the physical imagery that attaches to such a notion. Nothing actually *moved* in optics—I set aside the Newtonian fiction of "corpusles in flight"—until the invention of the dynamical wave theory of light (foreshadowed in 1678 by Huygens), where "rays" arise as "curves normal to surfaces of constant phase," and the "things" which literally move along "rays" are no more "physical" than mere points of intersection!

In mechanics, on the other hand, we confront the "real" motion of (idealized) "real things": mass points. I return now to the mechanical discussion where we left in on p. 4, asking ...

4. Does there exist a mechanical analog of Fermat's Principle? Such a theory, if it existed, would refer presumably to the *geometry of the space curves* $\boldsymbol{x}(s)$ traced out (in time) by m. This *separately from properly dynamical matters*, which can be considered to reside in the structure of the function $s(t)$. We adopt, therefore, this non-standard point of view:

$$\begin{array}{l} \boldsymbol{X}(s) \quad \text{describes the “trajectory” of } m\text{, the mechanical analog of a “ray”} \\ \big\downarrow \quad \uparrow \\ \quad s(t) \quad \text{describes } \textit{motion along} \text{ that trajectory} \\ \big\downarrow \\ \boldsymbol{x}(t) \equiv \boldsymbol{X}(s(t)) \quad \text{is the central object of Newtonian mechanics} \end{array}$$

To get off the ground we must recall some aspects of the mathematical theory of space curves. Let $\boldsymbol{X}(s)$ serve, relative to a Cartesian frame, to provide

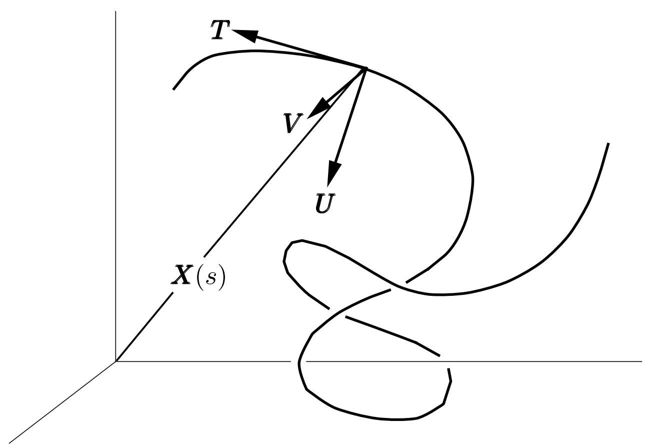

FIGURE 4: *Vectors natural to the description of a space curve* $\mathcal{C}$.

the s-parameterized description of such a curve $\mathcal{C}$. Then $\boldsymbol{T}(s) \equiv \mathring{\boldsymbol{X}}(s)$ describes the unit tangent to $\mathcal{C}$ at s. $\mathring{\boldsymbol{T}}(s)$ lies normal to $\boldsymbol{T}(s)$ and in the plane in which $\mathcal{C}$ is locally inscribed. The equation $\mathring{\boldsymbol{T}}(s) = \kappa(s)\boldsymbol{U}(s)$ serves (with $\boldsymbol{U}(s)$ a unit vector) to describe both the direction $\boldsymbol{U}(s)$ and the magnitude $\kappa(s)$ of the local *curvature* of $\mathcal{C}$. Assume $\kappa(s) \neq 0$ and define $\boldsymbol{V}(s) \equiv \boldsymbol{T}(s) \times \boldsymbol{U}(s)$ which serves to complete the construction of an orthonormal triad at each (non-straight) point s of $\mathcal{C}$. Elementary arguments lead to the conclusions that $\mathring{\boldsymbol{U}} = -\kappa\boldsymbol{T} - \tau\boldsymbol{V}$ and $\mathring{\boldsymbol{V}}(s) = \tau\boldsymbol{U}$, where $\tau(s)$ is the *torsion* of $\mathcal{C}$ at s. Briefly

$$\begin{pmatrix} \mathring{\boldsymbol{T}} \\ \mathring{\boldsymbol{U}} \\ \mathring{\boldsymbol{V}} \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & -\tau \\ 0 & \tau & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{T} \\ \boldsymbol{U} \\ \boldsymbol{V} \end{pmatrix}$$

which comprise the famous “Frenet-Serret formulae” (1847–1851).

Turning in this language to the subject matter of elementary kinematics, we have

$$\begin{aligned}\dot{\boldsymbol{x}} &= \dot{s}\,\mathring{\boldsymbol{X}}\\ &= \dot{s}\,\boldsymbol{T}\end{aligned}$$

$$\begin{aligned}\ddot{\boldsymbol{x}} &= \ddot{s}\,\boldsymbol{T} + \dot{s}^2\,\mathring{\boldsymbol{T}}\\ &= \ddot{s}\,\boldsymbol{T} + \underbrace{\dot{s}^2\kappa}\,\boldsymbol{U}\end{aligned}$$

$\kappa = 1/R$ with $R =$ "radius of curvature," so $\dot{s}^2\kappa$ is precisely the v^2/R familiar from the elementary theory of uniform circular motion

$$\begin{aligned}\dddot{\boldsymbol{x}} &= (\dddot{s} - \dot{s}^3\kappa^2)\,\boldsymbol{T} + (3\dot{s}\ddot{s}\kappa + \dot{s}^3\mathring{\kappa})\,\boldsymbol{U} - \dot{s}^3\kappa\tau\,\boldsymbol{V}\\ &\;\vdots\end{aligned}$$

which we now use to construct a "Newtonian theory of dynamical trajectories." To describe the conservative *motion* of m let us write

$$\ddot{\boldsymbol{x}}(t) = \boldsymbol{G}(\boldsymbol{x}(t))$$

with

$$\boldsymbol{G}(\boldsymbol{x}) \equiv -\tfrac{1}{m}\boldsymbol{\nabla}U(\boldsymbol{x}) = \tfrac{1}{2}\boldsymbol{\nabla}\Big\{\tfrac{2}{m}\big[E - U(\boldsymbol{x})\big]\Big\}$$

which we are now in position to formulate

$$\dot{s}^2\,\mathring{\boldsymbol{T}} + \ddot{s}\,\boldsymbol{T} = \boldsymbol{G} \tag{7}$$

All reference to the specifically *motional* aspects of the situation resides here in the factors $\dot{s}^2$ and $\ddot{s}$, which we now eliminate to obtain a "theory of trajectories" as a kind of residue. To that end, we note first that by energy conservation

$$\dot{s}^2 = \tfrac{2}{m}\big[E - U(\boldsymbol{x})\big]$$

Also by $\mathring{\boldsymbol{T}} \perp \boldsymbol{T}$ and $\boldsymbol{T}\cdot\boldsymbol{T} = 1$

$$\ddot{s} = \boldsymbol{T}\cdot\boldsymbol{G} = \tfrac{1}{2}\boldsymbol{T}\cdot\boldsymbol{\nabla}\Big\{\tfrac{2}{m}\big[E - U(\boldsymbol{x})\big]\Big\}$$

Partly to reduce notational clutter, but mainly to facilitate comparison with our optical experience, we agree to write

$$\tfrac{2}{m}\big[E - U(\boldsymbol{x})\big] \equiv v^2(\boldsymbol{x}) \equiv \left[\frac{c}{n(\boldsymbol{x};E)}\right]^2$$

where the c has been introduced from dimensional necessity but immediately drops away, and where it becomes natural to adopt the terminology

$$n(\boldsymbol{x};E) \equiv \text{ the “}\textit{mechanical}\text{ index of refraction”}$$

Returning with this information and in this notation to (7) we have

$$\tfrac{1}{n^2}\mathring{\boldsymbol{T}} + \tfrac{1}{2}\big(\boldsymbol{T}\cdot\boldsymbol{\nabla}\tfrac{1}{n^2}\big)\boldsymbol{T} = \underbrace{\tfrac{1}{2}\boldsymbol{\nabla}\tfrac{1}{n^2}}$$

But

$$= \tfrac{1}{n}\boldsymbol{\nabla}\tfrac{1}{n}$$

so

$$\tfrac{1}{n}\mathring{\boldsymbol{T}} + \big(\boldsymbol{T}\cdot\boldsymbol{\nabla}\tfrac{1}{n}\big)\boldsymbol{T} = \boldsymbol{\nabla}\tfrac{1}{n}$$

of which

$$\frac{d}{ds}\left[\frac{1}{n}\frac{d\boldsymbol{X}}{ds}\right] - \boldsymbol{\nabla}\frac{1}{n} = \boldsymbol{0} \tag{8}$$

is but a notational variant.

Equation (8) is the “trajectory equation” of Newtonian dynamics. It describes the “design of the tracks” along which m is permitted to move with conserved energy E. To describe the particle’s *progress along* such a track $\mathcal{C}_E$ we can appeal to the transit time formalism, writing

$$\begin{aligned} t(s) &\equiv \text{transit time } \boldsymbol{X}(0) \longrightarrow \boldsymbol{X}(s) \text{ along } \mathcal{C}_E \\ &= \int_0^s \frac{1}{\sqrt{\frac{2}{m}\big[E - U(\boldsymbol{X}(s'))\big]}}\,ds' = \tfrac{1}{c}\int_0^s n(\boldsymbol{X}(s');E)\,ds' \\ &\downarrow \\ s &= s(t) \quad \text{by functional inversion} \end{aligned}$$

We anticipate that there will be occasions when it is the intractability of the functional inversion that prevents our progressing from the trajectory to an explicit description of the motion—occasions, that is to say, when it is relatively easier to solve (8) than it is to solve the associated equations of motion.

The trajectory equation (8) provides the foundation of what might be called “time-independent Newtonian dynamics.” Interestingly, the phrase is much less familiar than the “time-independent Hamilton-Jacobi equation” and the “time-independent Schrödinger equation” which it calls instantly to mind. Nor are we speaking here of a merely terminological resonance; there exists a sense—which I hope to detail on some other occasion—in which the former subject lies at the theoretical base of the latter two.

5. Variational formulation of time-independent Newtonian mechanics. Equation (8)—the “trajectory equation”—is structurally identical to the “ray equation” (6), from which however it differs in one important respect, which can be symbolized

$$n(\boldsymbol{x})_{\text{optical}} \longrightarrow \frac{1}{n(\boldsymbol{x};E)_{\text{mechanical}}} \tag{9}$$

We have seen that the ray equation can be obtained by specialization $s \longleftarrow \lambda$ of the arbitrarily parameterized Euler-Lagrange equation (5) which issues from the variational principle $\delta\int n(\boldsymbol{x})\sqrt{\overset{\circ}{\boldsymbol{x}}\cdot\overset{\circ}{\boldsymbol{x}}}\,d\lambda = 0$. Similarly,

$$\delta\int\frac{1}{n(\boldsymbol{x};E)}\sqrt{\overset{\circ}{\boldsymbol{x}}\cdot\overset{\circ}{\boldsymbol{x}}}\,d\lambda = 0 \tag{10}$$

gives

$$\left\{\frac{d}{d\lambda}\frac{\partial}{\partial\overset{\circ}{\boldsymbol{x}}} - \frac{\partial}{\partial\boldsymbol{x}}\right\}\frac{1}{n(\boldsymbol{x};E)}\sqrt{\overset{\circ}{\boldsymbol{x}}\cdot\overset{\circ}{\boldsymbol{x}}} = \boldsymbol{0}$$

whence

$$\frac{d}{d\lambda}\left[\frac{1}{n}\frac{\overset{\circ}{\boldsymbol{x}}}{\sqrt{\overset{\circ}{\boldsymbol{x}}\cdot\overset{\circ}{\boldsymbol{x}}}}\right] - \sqrt{\overset{\circ}{\boldsymbol{x}}\cdot\overset{\circ}{\boldsymbol{x}}}\,\boldsymbol{\nabla}\frac{1}{n} = \boldsymbol{0} \tag{11}$$

from which by $s \longleftarrow \lambda$ we recover the trajectory equation (8).

Suppose were were to agree—at high risk of the confusion which I want here to dispel—to notate the variational principle (10) as follows:

$$\hat{\delta}\int\frac{1}{n}\,ds = 0 \tag{12}$$

where the $\hat{}$ is understood to signify that the variation is to be carried out subject to the constraint $\sqrt{\overset{\circ}{\boldsymbol{x}}\cdot\overset{\circ}{\boldsymbol{x}}} = 1$. Using

$$\frac{1}{n} = \sqrt{\frac{2}{mc^2}[E-U]} = \sqrt{\frac{2}{mc^2}}\cdot\sqrt{T} \quad \text{with } T = \text{kinetic energy}$$

we find that Newtonian trajectories of energy E have the property that they extremize "Jacobi's action functional"

$$\begin{aligned} A[\text{path}] &\equiv \sqrt{\tfrac{m}{2}}\int\sqrt{T}\,ds \\ &= \int T\,dt \quad \text{by } ds = \sqrt{\tfrac{2}{m}T}\,dt \end{aligned}$$

It is at this point that the standard literature becomes, by my reading, quite confusing. We have on the one hand

$$\textsc{hamilton's principle:}\quad \delta S = 0 \quad\text{with}\quad S[\text{path}] = \int(T-U)\,dt$$

and on the other hand

$$\textsc{jacobi's principle:}\quad \hat{\delta} A = 0 \quad\text{with}\quad A[\text{path}] = \int T\,dt \tag{13}$$

—*both* of which are known informally as the "Principle of Least Action," but the meanings of which are profoundly distinct. Hamilton's principle, as is well-known, has everything to do with the *temporal* aspects of dynamics: it gives us (in Lagrangian form) the *equations of motion*. Jacobi's principle, though deceptively notated to suggest otherwise, has in fact *nothing* to do

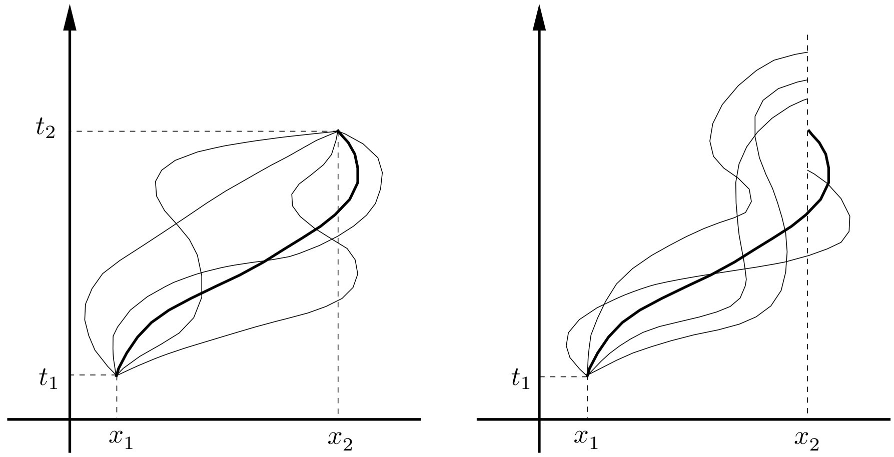

FIGURE 5: *Comparison of the variational processes fundamental to Hamilton's Principle (on the left) and the Jacobi's Principle of Least Action. In both cases the spatial endpoints are specified, but in Hamilton's Principle* transit time *is a constant of the variational process, while Jacobi's Principle requires the conserved* energy E *to be variationally constant.*

with temporal dynamics, but everything to do with the *geometrical figure of dynamical trajectories*; this becomes clear when one traces backwards the argument which led from (10) to (13). Sharpened understanding of the point here at issue follows at once from the observation that the integrand

$$\mathcal{A}_E(\boldsymbol{x}, \mathring{\boldsymbol{x}}) \equiv \frac{1}{n(\boldsymbol{x}; E)}\sqrt{\mathring{\boldsymbol{x}}\cdot\mathring{\boldsymbol{x}}}$$

in (10) is *homogeneous of degree one* in the variables $\mathring{\boldsymbol{x}} \equiv \frac{d}{d\lambda}\boldsymbol{x}$; under arbitrary reparameterization $\lambda = \lambda(\tau) \longleftarrow \tau$ we therefore have

$$\int \mathcal{A}_E(\boldsymbol{x}, \tfrac{d\boldsymbol{x}}{d\lambda})\,d\lambda = \int \tfrac{d\tau}{d\lambda}\mathcal{A}_E(\boldsymbol{x}, \tfrac{d\boldsymbol{x}}{d\tau})\cdot\tfrac{d\lambda}{d\tau}\,d\lambda = \int \mathcal{A}_E(\boldsymbol{x}, \tfrac{d\boldsymbol{x}}{d\tau})\,d\tau$$

according to which not only Jacobi's principle (10) but all of its consequences are *form-invariant with respect to arbitrary reparameterizations*. They therefore are, in particular, form-invariant with respect to arbitrary clock-regraduations $t \longrightarrow \tau = \tau(t)$, and so can have nothing to do with the specifically temporal aspects of mechanics.

We come thus to a conclusion which is, at least in the light of our optical experience, somewhat counterintuitive: the particle elects to pursue not the path $a \longrightarrow b$ which minimizes transit time, but the iso-energetic path which extremizes the Jacobi action $A[\text{path}]$; it pursues the "path of least action," a *geodesic with respect to the action metric*

$$d\sigma = a_E(\boldsymbol{x})ds$$

where $ds = \sqrt{(dx)^2 + (dy)^2 + (dz)^2}$ is Euclidean, and where

$$a_E(\boldsymbol{x}) \equiv \sqrt{\tfrac{2}{mc^2}\,[E - U(\boldsymbol{x})]} = \frac{1}{n(\boldsymbol{x}; E)}$$

With respect to the action metric the dynamical trajectory is "least curved," "straightest possible," generated by "parallel prolongation." The particle pursues its trajectory with Euclidean speed

$$\dot{s} = \sqrt{\tfrac{2}{mc^2}\,[E - U(\boldsymbol{x})]} = \frac{c}{n(\boldsymbol{x}; E)} = c \cdot a_E(\boldsymbol{x})$$

With respect to the action metric its "speed" is $\dot{\sigma} = a_E(\boldsymbol{x})\dot{s}$, and with respect to both metrics the speed is, in general, non-constant. With respect to the "reciprocal metric"

$$d\tilde{\sigma} = \frac{1}{a_E(\boldsymbol{x})}ds$$

speed *is* constant ($\frac{d}{dt}\tilde{\sigma} = c$), but the trajectory is non-geodesic.

6. Theoretical placement of the Principle of Least Action. It is my sense that the profoundly geometrical purport of Jacobi's principle is not widely appreciated, that physicists—even those writing about the subject (see, for example, §8.6 of Goldstein's 2$^{\text{nd}}$ edition, or §7.5 of his 1$^{\text{st}}$)—typically don't know quite what to make of Jacobi's principle, which they find it easy therefore simply to ignore. It is on these grounds that I understand the fact that the "Principle of Least Action" terminology is so often misapplied. For example, Richard Feynman gave to the dissertation (Princeton, May 1942) in which he first described what has come to be known as the "Feynman sum-over-paths formalism"; i.e., in which he first had occasion to write

$$K(x, t; x_0, 0) = \int e^{\frac{i}{\hbar}S[\text{path}]}\,\mathcal{D}[\text{paths}]$$

. . . the title *The Principle of Least Action in Quantum Mechanics*, though what he clearly had in mind was a quantum generalization of Hamilton's principle.[1] A paper which might more properly have worn Feynman's original title (but to

[1] When the work was finally published (Rev. Mod. Phys. **20**, 367 (1948)) it wore new title: "Space-time approach to non-relativistic quantum mechanics," where the first adjective refers to the fact that the "paths" in question are inscribed on spacetime, and the second adjective disabuses readers of any presumption that the theory has something therefore to do with relativity.

which he in fact gave a different title[2]) was written recently by Richard Crandall, who uses path-integral methods to study the time-independent object

$$G(x,x_0;E) = \frac{1}{i\hbar}\int_0^\infty K(x,t;x_0,0)\,e^{\frac{i}{\hbar}Et}\,dt$$

But the recent work which draws most explicitly upon the geometry of classical trajectories is that having to do with chaos—particularly quantum chaos.

E. T. Whittaker (see §§105–107 of his ANALYTICAL DYNAMICS) has drawn attention to the close kinship which links the Principle of Least Action to (for example) Gauss' "Principle of Least Constraint" and to the even less well known "Appell formalism." This whole corner of physics—of what I have chosen to call "geometrical mechanics"—remains much less studied than (in my view) it deserves to be. Research questions spring easily to mind, and much that is useful (not to say pretty) remains to be worked out. It is not so much my best scientific judgment as simple sentiment which has recommended to my attention the little nest of interrelated topics to which I now turn.

7. Comparison with the mechanics of constrained free motion. Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N$ be the inertial Cartesian coordinates of a system of masses $m_1, m_2, \ldots, m_N$ which move subject to K holonomic constraints

$$\varphi_k(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N) = 0 \quad : \quad k = 1, 2, \ldots, K$$

and let $q^1, q^2, \ldots, q^n$ refer to some coordinatization of the n-dimensional space C_n of points which conform to those imposed constraints. Evidently C_n—which is effectively an $(n = 3N - K)$-dimensional surface inscribed in $R_3 \times R_3 \times \ldots \times R_3$—is a Riemannian manifold with "mass metric" $m_{ij}(q)$ inherited ultimately from the Euclidean structure of intertial 3-space. System motion traces a curve $q(t)$ in the "configuration space" C_n, and the "dynamical problem" is to describe that t-parameterized curve. This is precisely the problem which stimulated the invention of Lagrangian dynamics. One writes

$$\left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k} - \frac{\partial}{\partial q^k}\right\} L = 0$$

$$L \equiv T(q,\dot{q}) - U(q)$$

In the absence of impressed forces—i.e., if the only forces experienced by the particles are forces of constraint—then

$$U(q) \equiv 0$$

and we have

$$L = \tfrac{1}{2} m_{ij}(q)\dot{q}^i\dot{q}^j$$

---

[2] See "Combinatorial approach to Feynman path integration," J. Phys. A: Math. Gen. **26**, 3627 (1993) and papers there cited.

Now some remarks intended to expose more clearly some geometrical aspects of the situation:

Generally, L and $\mathcal{L} \equiv L^p$ give rise to distinct and inequivalent equations of motion. For we have

$$\left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k} - \frac{\partial}{\partial q^k}\right\}L^p = pL^{p-1}\left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k} - \frac{\partial}{\partial q^k}\right\}L + p(p-1)L^{p-2}\dot{L}\frac{\partial L}{\partial \dot{q}^k}$$

Clearly

$$\left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k} - \frac{\partial}{\partial q^k}\right\}L = 0 \quad \Longrightarrow \quad \left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k} - \frac{\partial}{\partial q^k}\right\}L^p = 0$$

only if it is, for some reason, automatic that $p(p-1)L^{p-2} \cdot \dot{L} \cdot (\partial L/\partial \dot{q}^k) = 0$

THEOREM: If $L(q, \dot{q})$ is t-independent and homogeneous of degree $n \neq 1$ in $\dot{q}$ then necessarily $p(p-1)L^{p-2} \cdot \dot{L} \cdot (\partial L/\partial \dot{q}^k) = 0$.

The proof is elementary: from $\partial L/\partial t = 0$ it follows that $\dot{J} = 0$, where

$$J = \sum \frac{\partial L}{\partial \dot{q}^k}\dot{q}^k - L$$

is "Jacobi's intergral." By Euler's theorem

$$\sum \frac{\partial L}{\partial \dot{q}^k}\dot{q}^k = nL$$

so we have $J = (n-1)L$ and (since by assumption $n \neq 1$) $\dot{J} = 0 \Rightarrow \dot{L} = 0$, which is sufficient to establish the result claimed. ■

The application of immediate interest (set $p = \frac{1}{2}$ and note that kinetic energy T in homogeneous of degree 2 in the generalized velocities) reads

$$\left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k} - \frac{\partial}{\partial q^k}\right\}T = 0 \quad \Longrightarrow \quad \left\{\frac{d}{dt}\frac{\partial}{\partial \dot{q}^k} - \frac{\partial}{\partial q^k}\right\}\sqrt{T} = 0$$

But the latter can, by Hamilton's principle, be formulated $\delta \int \sqrt{T}\, dt = 0$ or again

$$\delta \int ds = 0 \quad \text{where } ds = \text{differential length with respect to the mass metric}$$

Evidently $q(t)$ traces with constant speed[3] a geodesic in C_n. If $N = 1$ (which is to say: for single particle systems) all this is directly evident to the eye, and illustrated in the following figure. Note that adjustment of the arrival time

[3] This not at all surprisingly: *forces of constraint do no work!*

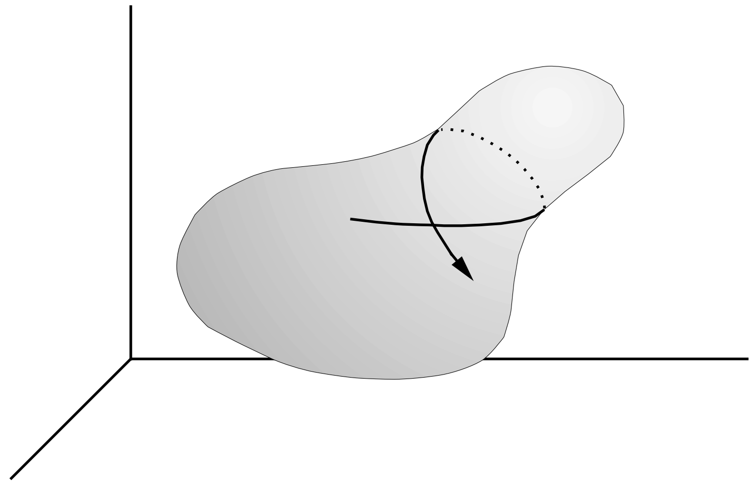

FIGURE 6: *Constrained free motion of a single particle, which pursues a geodesic with constant speed.*

entails adjustment of the speed/energy of the mass point(s), but not of the geodesic trajectory: the metric is now (because E enters as a detachable factor) effectively E-independent. While I have made use here of the methodology of Lagrange and Hamilton, identical conclusions would have been obtained—and more directly—had I worked from Jacobi's principle (12) with $U = 0$.

8. Enter: Heinrich Hertz. . . . who was born (in Hamburg) on 22 February 1857, and who died—aged 36—just 100 years ago last New Year's Day. Hertz went in 1877 to Munich to study engineering, but discovered almost immediately that it was physics to which he was most strongly attracted. Accordingly, he gave the winter of 1877/78 to private study of the classics (Laplace, Lagrange, . . .) and in 1878 removed himself to Berlin, where he studied under Kirchhoff and H. von Helmholtz. There he plunged immediately into experimental research. In 1880 his study of "Kinetic Energy of Electricity in Motion" won a prize, and he received his PhD for a dissertation treating "Induction in Rotating Spheres." In 1883, after three years of work as Helmholz' assistant, he became *privatdozent* in Kiel and began his study of Maxwell's electromagnetic theory. In 1885 he moved to Karlsruhe, where he married (the daughter of a colleague) and carried out the experiments for which he became instantly famous. In 1879 Helmholtz had directed his attention to a prize being offered by the Berlin Academy of Sciences for "experimental establishment of some relation betwen electromagnetic forces and the dielectric polarization of insulators," but Hertz could at the time think of no way to attack the problem. In 1887 he did think of a way, and established the physical reality of "displacement current." He generated, detected and studied basic properties of electromagnetic radiation in 1888. By 1889 his fame was so great that he was called to Bonn, to succeed Rudolph Clausius. While in Bonn a colleague—seven years his junior—was

Hermann Minkowski, who on one occasion wrote pleadingly to his friend David Hilbert that "I must get out of here or I'll become a physicist! The only person worth talking to here is Hertz."

Hertz was a first class experimentalist, but by no means was he only an experimentalist; in the extraordinary range of his creative interests he was reminiscent of his principal mentor, the great Hermann von Helmholtz.[4] It was therefore quite in character for him, while in Bonn (where he developed the first symptoms of the cancer of the jaw that ultimately killed him, and found himself therefore unable to pursue an experimental program), to turn his attention—particularly his critical attention—to theoretical mechanics and (very much in the tradition of Herschel, Thompson & Tait, Maxwell, Clifford, Mach, Poincaré, ...) the philosophy of science.[5]

During the course of this work, Hertz independently reinvented Jacobi's Principle of Least Action, and also Gauss' Principle of Least Constraint (which Hertz called the "Principle of Least Curvature"). But for Hertz these were of interest only as steps toward a grander objective: for philosophical reasons which I could not begin to summarize,[6] Hertz held it to be a defect of Newtonian mechanics that *two* notions—"inertia" and "force"—are required to account for the motion of mass points. His objective was to construct a formalism within which it could be held that *all seeming "forces" arise actually as "forces of constraint"* (therefore as artifacts off inertia), and to that end he was prepared to *adjoin to the familiar space variables x, y and x some variables* $\alpha_1, \alpha_2, \ldots, \alpha_n$ *which refer to "hidden dimensions" of physical space.*[7]

That a Herzian "forceless mechanics" *exists* was first brought to my own attention by A. O. Barut when I was a Reed College student (1954/55), and he my teacher. But how it works remained obscure to me (Hertz was a turgid writer) until—quite recently, and quite by accident—I came upon a copy of

[4] Helmholtz was 36 years older than Hertz, and died in the September of 1894, eight months after Hertz. One of his last professional acts was to prepare Hertz' papers for posthumous publication and to write his obituary.

[5] The philosophical work of these thinkers actually affected the way their fellow scientists went about their work (recall the inspiration Faraday drew from Herschel, Einstein from Mach); moreover, it was accorded respect by professional philosophers of the day, and pondered by the generality of active intellectuals. Allan Janik & Stephen Toulmin, in *Wittgenstein's Vienna* (1973), provide a good account of the vanished tradition to which I refer. They mention, by the way, that Wittgenstein was always frank to acknowledge his special indebtedness to the thought of Hertz.

[6] See Hertz' *Principles of Mechanics*, published posthumously (with a foreword by Helmholtz) in 1899, and in English translation by Dover in 1956.

[7] The philosophical thrust of Hertz' program brings to mind the influence which, fifteen years later, the Principle of Equivalence exerted upon the thought of Einstein: "If m_{inertial} and $m_{\text{gravitational}}$ are universally proportional, it must be because, from some sufficiently radical point of view, they are *the same thing*.

Barut's *Geometry and Physics: Non-Newtonian Forms of Dynamics* (1989). It is to Barut that I owe the essentials of what now follows.

9. Hertz' "forceless mechanics." Hertz' idea, in its simplest essentials, can be understood if we start in N-dimensional Euclidean space ($N > 4$) with the "free particle" system

$$\mathcal{L} = \tfrac{1}{2}m(\dot{X}_1^2 + \dot{X}_2^2 + \cdots + \dot{X}_N^2) \tag{13}$$

and impose such holonomic constraints

$$\left.\begin{aligned} \varPhi_1(X) &= 0 \\ &\vdots \\ \varPhi_{N-4}(X) &= 0 \end{aligned}\right\} \tag{14}$$

that the configuration space becomes 4-dimensional, with the "mass metric" (or "Hertz metric") implicit in

$$L = \tfrac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) + \frac{1}{U(x)}\dot{\alpha}^2 \tag{15}$$

Here x_1, x_2, x_3 and α are the names we have given to what are, in effect, "generalized coordinates," and (it is interesting to note) α has necessarily the dimensionality of "action." The resulting Lagrange equations read

$$m\ddot{\boldsymbol{x}} - \dot{\alpha}^2\boldsymbol{\nabla}\tfrac{1}{U} = \boldsymbol{0} \quad \text{and} \quad \tfrac{d}{dt}\big[\tfrac{1}{U}\dot{\alpha}\big] = 0 \tag{16}$$

The latter gives $\dot{\alpha} = kU$ where k is some dimensionless constant. Returning with this information to the former equation, we have

$$\begin{aligned} m\ddot{\boldsymbol{x}} &= k^2U^2\boldsymbol{\nabla}\tfrac{1}{U} \\ &= -k^2\boldsymbol{\nabla}U(\boldsymbol{x}) \end{aligned} \tag{17}$$

which (i) contains no reference to the variable α except covertly (through k) and (ii) looks Newtonian except for the unfamiliar constant k, which we can either *interpret to be a parameter controlling the strength of the potential* or (alternatively) *eliminate by rescaling the time variable*. From the "other equation of motion" we obtain $k = \dot{\alpha}(0)/U(\boldsymbol{0})$, which shows the specific value of k to be *implicit in the initial data*; in Newtonian mechanics we are traditionally content to insert initial data after the fact, by hand, but in Hertz' forceless mechanics it enters into the essential design of the dynamical equation.

Similarly

$$L = \tfrac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) + \frac{1}{U_1(\boldsymbol{x})}\dot{\alpha}_1^2 + \frac{1}{U_2(\boldsymbol{x})}\dot{\alpha}_2^2$$

gives

$$\begin{aligned} m\ddot{\boldsymbol{x}} &= -k_1^2\boldsymbol{\nabla}U_1(\boldsymbol{x}) - k_2^2\boldsymbol{\nabla}U_2(\boldsymbol{x}) \\ &= -k_1^2\boldsymbol{\nabla}(U_1 + \lambda U_2) \quad \text{with} \quad \lambda \equiv (k_2/k_1)^2 \end{aligned}$$

But so also does

$$L = \tfrac{1}{2}m(\dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2) + \frac{1}{U_1(\boldsymbol{x}) + \lambda U_2(\boldsymbol{x})}\dot{\alpha}^2$$

So a certain *non-uniqueness* attaches to the Hertz formalism, lending force to claims that Hertz made not (as he had intended) a physical discovery, but a discovery of merely formal purport. The formalism does, however, have at least the attractive property that it is deeply geometrical in spirit. In $\boldsymbol{x}$-space the

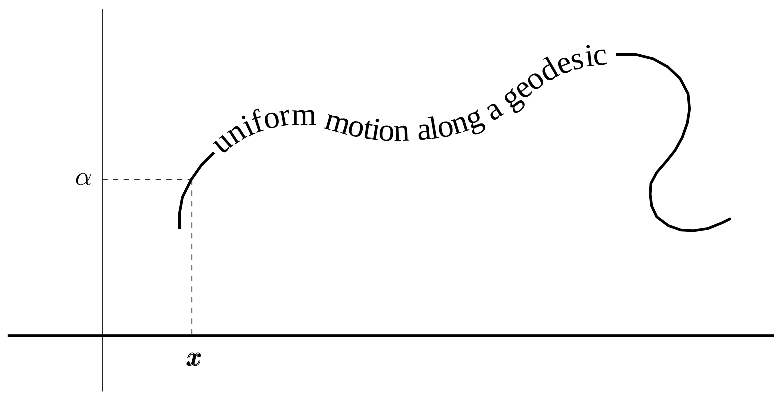

FIGURE 7: *With respect to a suitably designed "Hertz metric" in an appropriate hyperspace, any conservative force can be represented as a "force of constraint," and all motion becomes "constrained free motion."*

motion of m is non-uniform, and geodesic only with respect to the E-dependent "action metric." But in $(\alpha, \boldsymbol{x})$-space the motion is *uniform geodesic with respect to the Hertz metric*; it is, in short, *free on a curved surface in N-space*, and *the only forces experienced by m are forces of constraint*.

10. The embedding problem. In progressing from (13) to (17) I (in the interests of expository swiftness) tacitly assumed a certain sequence of procedures to be *possible*, and have now to consider whether that assumption is in fact tenable.[8]

Hertz' equations of motion (16) can be obtained from a hyper-dimensional variant of Hamilton's principle

$$\delta \int L\, dt = 0$$

where the hyper-Lagrangian L is (see again (15)) homogeneous of degree two in the generalized velocities $\dot{\boldsymbol{x}}$ and $\dot{\alpha}$. Drawing now upon the THEOREM of p. 14, we obtain the equivalent statement

$$\delta \int \sqrt{L}\, dt = 0 \tag{18}$$

[8] It turns out *not* to be, or at least not quite.

where $\sqrt{L}\,dt = \sqrt{\frac{m}{2}}\,ds$ and where

$$ds^2 = dx^2 + dy^2 + dz^2 + A(\boldsymbol{x})d\alpha^2 \qquad A(\boldsymbol{x}) \equiv \tfrac{2}{m}\cdot\tfrac{1}{U(\boldsymbol{x})} \tag{19}$$

serves to define the "Hertz metric." The Hertz metric ascribes metric structure to Hertz' configuration space C_4. Our problem—the "embedding problem"—is to display C_4 as a 4-dimensional surface in an enveloping flat E_N, a surface so constructed that the Hertz metric is *induced* by the (simpler) metric structure of E_N.

Embedding problems (like most "inverse problems") are hard, and so in particular I found this one to be. It was, however, cracked "a long time ago" by one of Barut's students (and a wonderfully clever student she must have been!). I work now from unpublished material Barut was kind enough to share with me.

Let x, y, z, u, v, w coordinatize an E_6 with the *psuedo*-Euclidean metric structure defined

$$ds^2 = dx^2 + dy^2 + dz^2 + du^2 + dv^2 - dw^2$$

Let constraints

$$\left.\begin{aligned} \Phi_1(\boldsymbol{x},u,v,w) &\equiv v - \tfrac{1}{2}\sqrt{A(\boldsymbol{x})}\Big[1 - \frac{u^2}{A(\boldsymbol{x})}\Big] = 0 \\ \Phi_2(\boldsymbol{x},u,v,w) &\equiv w - \tfrac{1}{2}\sqrt{A(\boldsymbol{x})}\Big[1 + \frac{u^2}{A(\boldsymbol{x})}\Big] = 0 \end{aligned}\right\} \tag{20}$$

serve to define a surface Σ_4 in E_6. Easily $u^2 + v^2 - w^2 = 0$ so the constraints (20) serve to erect at each $\boldsymbol{x}$-point a certain (u, v, w)-cone, and to inscribe on that cone a certain curve $\mathcal{C}$. The equations

$$\begin{aligned} u &= \sqrt{A(\boldsymbol{x})}\,\alpha \\ v &= \tfrac{1}{2}\sqrt{A(\boldsymbol{x})}\,(1-\alpha^2) \\ w &= \tfrac{1}{2}\sqrt{A(\boldsymbol{x})}\,(1+\alpha^2) \end{aligned}$$

provide a fairly natural parameterization of that curve. By straightforward calculation we now find

$$\begin{aligned} du &= \sqrt{A}\;\;d\alpha + \tfrac{1}{2\sqrt{A}}\alpha\boldsymbol{\nabla}A\cdot d\boldsymbol{x} \\ dv &= -\sqrt{A}\alpha\,d\alpha + \tfrac{1}{4\sqrt{A}}(1-\alpha^2)\boldsymbol{\nabla}A\cdot d\boldsymbol{x} \\ dw &= +\sqrt{A}\alpha\,d\alpha + \tfrac{1}{4\sqrt{A}}(1+\alpha^2)\boldsymbol{\nabla}A\cdot d\boldsymbol{x} \end{aligned}$$

giving (after further calculation)

$$du^2 + dv^2 - dw^2 = A(\boldsymbol{x})\,d\alpha^2$$

So
$$ds^2 = dx^2 + dy^2 + dz^2 + A(\boldsymbol{x})\,d\alpha^2$$
and the desired embedding has been achieved. But by a hat trick, which has contributed disappointingly little to general understanding of the "art of embedding."

In a more complete account of this topic it would become important to notice that $A \longrightarrow \infty$ at points where $U(\boldsymbol{x}) = 0$, and that $\sqrt{A}$ becomes imaginary at points where $U(\boldsymbol{x}) < 0$. Were we to insist (on what grounds?) that E_6 be a *real* psuedo-Euclidean space then we would become obligated to write $\sqrt{-A}$ where formerly we wrote $\sqrt{A}$, and to assign to E_6 this altered metric structure:

$$ds^2 = dx^2 + dy^2 + dz^2 - du^2 - dv^2 + dw^2$$

It is, however, not immediately evident what sense we are to make of "an embedding space which *changes from point to point*," depending on the sign of $U(\boldsymbol{x})$. Relatedly, we would, in a more complete discussion, want to describe how the formalism responds to unphysical adjustments of the type $U \to U+$constant.

Summary, and a glance ahead. Study of the transit time concept in a multi-dimensional context led (explicitly in the conversation from which this work derives, but only implicitly in the actual text) to the formulation of a

> CONJECTURE: When moving $a \to b$ in a conservative force field, a particle selects the path which minimizes transit time T[path].

which proved the more interesting for being false! We were led thus to the realization that Fermat's Principle of Least Time and the Principle of Least Action (Jacobi/Hertz) are—though both refer not to motion but to the *geometry of physical space curves* (not to the "progress of the train" but to the "design of the track")—not formal analogs of one another, but formal "reciprocals," with consequences which (or so I have come to suspect) show up ultimately as the reciprocity between phase velocity and group velocity.

En route to that understanding we gained new appreciation of the physical utility of the classical theory (Frenet-Serret) of space curves. And, having arrived, we found ourselves in position to understand why the "optical ray equation" and the "Newtonian trajectory equation"—both of which are less well known than they deserve to be—are, while similar, yet not identical.

In extracting Jacobi's Principle (not, as is standard in the literature, from Hamilton's Principle but) directly from Newton's Second Law we have (in my view) *simplified the derivation*[9] and *more clearly exposed the geometrical meaning and latent power* of the Principle of Least Action. Which, as we have

---

[9] Along the way we had occasion to develop a little theorem according to which, under conditions fairly standard to the "geometrization of dynamics," $\delta\int L\,dt = 0 \Leftrightarrow \delta\int\sqrt{L}\,dt \sim \delta\int ds = 0$. The same theorem, run backwards, permits one, in connection with the Riemannian diffential geometry of geodesics, to get rid of the square root which enters awkwardly into the definition of ds.

been led to emphasize, refers to certain *time-independent* aspects of classical motion. We are now positioned, therefore, to review corresponding aspects of time-independent quantum mechanics . . . which on some future occasion I propose to do.

Theorems relating to the convergence/divergence of geodesic trajectories were developed already by Jacobi himself,[10] whose language anticipates in some respects that of modern chaos theory. Periodic systems give rise to *closed* trajectories. It seems likely that new insight into the classical "theory of adiabatic invariants"—as well (relatedly) to the theory of "geometric phase" (Berry's phase)—can be extracted from our now sharpened understanding of Jacobi's Principle. And according to Whittaker (§§105 & 107) we stand now in position to gain sharpened insight also into the construction of Gauss' Principle of Least Constraint and of the Appell formalism.

Finally, we were brought to a better appreciation of the mind and accomplishment of Heinrich Hertz who, though universally reknowned for his experimental finesse, was clearly a theorist of exceptional imagination and power—remarkable not least for the fact that he drew much of his creative impulse from a *philosophical* train of thought. His work looks back to the work of Jacobi/Gauss, and forward to the work of Einstein (∼1915), Kaluza (1921) and Klein (1926). He emerges as a central—if too frequently overlooked—figure in the history of the progressive "geometrization of physics," and I find it easy to suppose that was a seed planted conversationally by Hertz which permitted Minkowski (∼ 13 years later) to leap to the realization that Einstein had, in effect, "geometized spacetime." Hertz invented a context within which it became possible to *think of "force" as an artifact of broken symmetry*. But on a variety of mainly formal grounds (the theory is, for example, at many points arbitrary, and provides no evident principles for selecting among the alternatives) I myself—in the company, I gather, of most other physicist—am disinclined to ascribe any significant degree of physically plausibility to Hertz' "forceless mechanics."[11] Certainly it is in any event undeniable that in recent years "hyperdimensionalization" has become a thriving industry, and Hertz was (so far as I am aware) the first to venture down that fertile road. I lift my hat to the memory of the man.

ADDENDUM. A wonderful book has appeared, which bears directly upon the "embedding problem" encountered in §10. I refer to Sylvia Nasar's *A Beautiful Mind: A Biography of John Nash* (1998). Chapter 20 provides an account of the circumstances which led to the production of Nash's monumental "The imbedding problem for Riemannian manifolds," Ann. Math **63**, 1956.

---

[10] See E. T. Whittaker's §103.

[11] Hertz' theory does, on the other hand, have (it seems to me) value as a kind of theoretical laboratory, a place where one can study formal points which may acquire physical interest in other connections; I would, in particular, like to know whether (and how) "forceless mechanics" admits of quantization.