CHAPTER 3

A LITTLE HISTORY

We saw that eq. (2.2.8) and, indeed most of this thesis, is based on the variational principle, eq.(2.1.16), which constitutes the basis of the analogy to geometrical optics and to classical mechanics. To clarify the physical significance of eq. (2.1.16), it will help to place this variational principle in the historical context of the optical-mechanical analogy. This is done in Sec.3.1. The other section (Sec.3.2) displays the relationship between the "F=ma" optics and Hamilton's procedure.

3.1. Historical background: a reapprisal of events

The history of the relationship between Fermat's principle and Maupertuis' principle spans the period from the seventeenth to the twentieth century and includes the two-and-a-half century debate over the nature of light and the rise of quantum mechanics. It is a story rich in irony and reversals of fortune. The story begins with Descartes' derivation of the law of refraction from a particle model for light.[48] Descartes likened a particle of light to a tennis ball that suffers an impulsive blow, in the direction of the normal to the boundary between two media as it passes over that boundary. The component of the velocity parallel to the boundary is therefore the same in both media, while the normal component changes. The constancy of the ratio of the sines of the angles of incidence and refraction immediately follows. Because light bends closer to the normal while passing from air to water, it follows that light must travel more rapidly in the water. Descartes was immediately imbroiled in controversy. For, in about same time that Descartes was applying his mechanical philosophy to light, the constancy of the ratio of the sines was discovered experimentally by Snell. Friends of Snell accused Descartes of plagiarism.

Twenty five years later, Pierre de Fermat succeeded in connecting the law of refraction with an extremum principle. Fermat proposed, in private correspondence, that light travels from a point in one medium to a point in another by the path requiring the least time.[49] From this Fermat was able to obtain the constancy of the ratio of the sines of the angles of incidence and refraction. but if light obeyed Fermat's principle, it must travel more rapidly in the air. Fermat was attacked by the followers of Descartes for advocating a principle which implies that light travels more rapidly in air than in water when the master had taught that just the opposite was true. Worse, Fermat appeared to be reintroducing teleological arguments and final causes into natural philosophy[50]. Fermat replied bitterly that he would leave to Mr. Descartes the glory of having explained the refraction of light and would content himself with offering an abstract mathematical proposition, without asserting that it applied to light.

In 1690, Christiaan Huygens published his *Treatise on Light*, in which he developed a wave theory of light, and applied it to rectilinear propagation, reflection, refraction and double refraction. Huygens also proved that, if light were a wave, it would obey Fermat's principle of least time. Thus, at the close of the seventeenth century, Fermat's principle was attached to the wave theory.[51]

Newton, in his *Opticks* of 1704, advocated a corpuscular theory of light--although his corpuscles had to be endowed with a vibratory character to explain the inflections of light (or, as we would say, diffraction and interference phenomena). The corpuscularity of light was practically the only proposition of physics on which Newtonians and Cartesians could agree. It is not surprising that, under the combined influence of Newton and Descartes, the wave theory practically disappeared. And Fermat's principle of least time fell into disrepute: hardly anyone believed in it. It is not to be found in most of the eighteenth-century textbooks, for example. There were in the eighteenth century always a few advocates of the wave theory, among whom Euler was the most prominent. But the overwhelming majority of physical thinkers believed that light is a stream of particles and that, as both Descartes and Newton had shown, it must therefore travel more quickly in water than in air.

In 1744, Maupertuis announced a minimization principle intended to encompass both mechanics and optics: the principle of least action.[52] Maupertuis called *action* the product of mass and distance. His definition of action was vague and was applied inconsistently, but he succeeded in deducing from his principle the rules governing elastic and inelastic collisions of particles. More to the point for our story, he also applied his principle to light and managed to derive the law of refraction. Maupertuis' principle was consistent with the view of light as a particle and the opinion of Descartes and Newton that the speed of light was greater in denser media. With this fact of nature re-established, Maupertuis exclaimed, "the whole edifice that Fermat had constructed is destroyed: light, when it traverses different media, goes neither by the shortest route, nor by the route of least time .... "[53]. Maupertuis also pointed out that for the cases of reflection and straight-line propagation --cases in which the speed of light does not change--the path of least time is the same as the path of least action, which, according to Maupertuis, explained how Fermat had gone wrong.

Thus, by the middle of the eighteenth century, there were two competing variational principles, least time and least action, each of which could serve as a basis for geometrical optics. One of these was associated with the wave theory and one with the particle theory. The *shape* of a ray in a given situation was predicted to be exactly the same by the two theories. The theories differed only in the progress in time of light along the ray; but, of course, this could not be measured.

The story told so far is the prehistory of the optical-mechanical analogy. A major issue at stake was the nature of light. Proponents of the particle theory were not making an analogy between geometrical optics and mechanics but rather *claiming* that they were one and the same. At the close of the eighteenth century, majority opinion held that one variational principle--that of Maupertuis--covered both. The minority school of wave theorists recognized two principles--those of Fermat and Maupertuis--which applied to different domains of physics.

The principle of least action *as it applies to mechanics* was elaborated and clarified by Euler, and especially by Lagrange. Lagrange formulated the least action principle more precisely, developed a calculus of variations for exploring its consequences and demonstrated that it could serve as a foundation for mechanics, a welcome alternative to the philosophically disturbing force-based physics of Newton. Among other things, Lagrange showed that principle of least action and conservation of energy were together equivalent to Newton's law of motion (F = ma)[54]. This also served to strip Maupertuis' principle of its metaphysical significance. For Maupertuis, the fact that both light and particles follow the paths of least action had been a sure sign of the wisdom of the Creator. For Lagrange, the principle of least action was just an alternative foundation for mechanics.

The principle of "least action" (really, least time) *as it applies to optics* was explored by Laplace,[55] but most notably by William Rowan Hamilton, who was largely responsible for the creation of the optical-mechanical analogy.[56] Hamilton's early work in mathematical physics was devoted to geometrical optics. When Hamilton began this work, in the 1820's, French physicists were already abondoning the particle theory of light, as a result of the success of Fresnel's mathematical wave theory. In England and Ireland (Hamilton's native land), where Newton's influence was stronger, the conservation required another decade or so. Thus, in his first papers on systems of rays (1824-1830), Hamilton always called the integral $\int n\,dl$ the *action*, in analogy with the action integral $\int v\,dl$ of mechanics. The analogy would be most straightforward for the particle theory, in which *n* is directly proportional to v. However, Hamilton did not necessarily subscribe to the particle theory. In his first paper, he refrained from stating a position on the nature of light and wrote that he used the term *action*, "intending only to express a remarkable analogy, and not assuming any hypothesis about the nature of light".[57] For the integrand of the "action" integral (essentially the index of refraction), Hamilton sometimes used m, sometimes μ, and sometimes v. The use of v for the refractive index might seem to imply acceptance of the particle theory, but this was not the case.

Indeed, by the time (1832) of his "Third Supplement", Hamilton had definitely adopted the wave theory. This is apparent in the softening of his vocabulary; for he now referred to a "law of least action, or of swiftest propagation", i.e., he began to adopt the language of the wave theory and to use it side-by-side with, and as an alternative to, the language of the particle theory.[58] The integrand v he called the *medium-function* and characterized it as a "a molecular velocity or an undulatory slowness". In the "Third Supplement", moreover, Hamilton applied his methods directly to the wave theory of Fresnel. The most spectacular result was Hamilton's theoretical prediction of the phenomenon of conical refraction, which was observed shortly afterward by Lloyd.

The essence of Hamilton's geometrical optics is his *characteristic function*

$$V(a,\ b,\ c,\ x,\ y,\ z) \equiv \int_{a,\ b,\ c}^{x,\ y,z} n\,dl \qquad (3.1.1)$$

which is nothing other than Hamilton's "action" integral or, (as we would say) the optical path length between point (a, b, c) and point (x, y, z) on the same ray. Hamilton showed that V satisfies the partial differential equation

$$\left|\frac{\partial V}{\partial x}\right|^2 + \left|\frac{\partial V}{\partial y}\right|^2 + \left|\frac{\partial V}{\partial z}\right|^2 = n^2(x,\ y,\ z). \qquad (3.1.2)$$

If this differential equation can be solved, knowledge of V constitutes a complete knowledge of the optical system. In particular, the direction of the ray at point can be found by differentiation of V. Hamilton's characteristic function V is today usually called the *eikonal*, a term introduced by H. Bruns in 1895. eq.(3.1.2) is called the *eikonal equation*, and is one of the fundamental equations of modern geometrical optics.[59] The surfaces V = constant are surfaces of equal travel time--the wavefronts.

In the same way, a central feature of Hamilton's mechanics is the characteristic function

$$V(a,\ b,\ c,\ x,\ y,\ z) \equiv \int_{a,\ b,\ c}^{x,\ y,z} v\,dl, \tag{3.1.3}$$

which is the action for a particle moving between points (a, b, c) and (x, y, z). Hamilton showed that, for a unit mass, V satisfies the partial differential equation

$$\left|\frac{\partial V}{\partial x}\right|^2 + \left|\frac{\partial V}{\partial y}\right|^2 + \left|\frac{\partial V}{\partial z}\right|^2 = v^2(x,\ y,\ z) = 2(E - U) \tag{3.1.4}$$

where U is a potential energy function and E is the energy, which is constant on a given trajectory. If this differential equation can be solved, the characteristic function V constitutes a complete description of the motion of the particle. For example, the velocity components of the particle at (x, y, z) can be found by differentiation.[60] Thus, in Hamilton's formulation of mechanics, it is the characteristic function that embodies the optical-mechanical analogy. In mechanics, the surfaces V = constant are surfaces of constant action, analogous to the surfaces of constant phase or travel time in optics. The particle trajectories correspond to the optical rays. Hamilton's paper on the optical-mechanical analogy, "On a General Method of Expressing the Paths of Light, and of Planets, by the Coefficients of a Characteristic Function", was published, obscurely, in the *Dublin University Review* for October, 1833.[61]

Hamilton has struggled long and hard with geometrical optics. The development of general theory of rays occupied him from 1824 to 1832. By contrast the works on mechanics came quickly and easily in the years 1833-34, for it was largely but an extension of the optical theory. It is one of the many ironies of this story that there exists today a Hamiltonian system of dynamics because Hamilton happened to develop his geometrical optics at a time when the debate over particle and wave theories of light had not yet been resolved.[62]

Hamilton's papers on dynamics won him lasting recognition. His methods in dynamics was further developed by Jacobi and remained a part of the main stream of nineteenth-century theoretical physics. In contrast, Hamilton's optical papers were read and understood by very few and soon dropped almost completely out of sight. There were several reasons for this. Hamilton's Irish national sentiment had led him to publish his optical papers in an obscure and little-circulated journal, the *Transactions* of the Royal Irish Academy, even though his uncle had warned him that this would be little better than committing them to a tomb.[63] (The papers on mechanics appeared in the *Philosophical Transactions* of the Royal Society of London). Perhaps an even greater obstacle was Hamilton's difficult style. The papers are very densely written. Moreover, Hamilton strove constantly for higher abstraction and greater generality, and rarely concerned himself with practical applications. He created a general system of geometrical optics, but failed to teach his prospective readers how to use it.

It is a further irony of this story that Bruns, the modern rediscoverer of Hamilton's method in optics, had seen Hamilton's optical papers. Bruns essentially worked backward from Hamilton's theory of characteristic function in mechanics to reach an optical analogy--the reverse of the route that Hamilton had traveled some seventy years before.[64]

By the 1830's, the wave theory had triumphed. Fermat was vindicated at last, and one could confidently express Fermat's principle in the form

$$\delta\int dl/v = 0. \qquad \text{(optics)} \qquad (3.1.5)$$

Now there were two well-established variational principles, eq.(3.1.5) for light and Maupertuis' principle for particles:

$$\delta\int v\,dl = 0. \qquad \text{(mechanics)} \qquad (3.1.6)$$

The fact that they were so similar in form was only a curiosity, and Hamilton's optical-mechanical analogy was regarded as an elaborate and sterile oddity, until the development of quantum theory in the first few decades of our own century.

The action integral of eq.(3.1.6) played a major role in the old quantum theory of Bohr and Sommerfeld. But it was Louis de Broglie who first brought the principles of Fermat and Maupertuis together and showed that they were one. de Broglie argued that atoms displayed effects involving integer numbers (as in the Balmer formula for the frequencies of light emitted by the hydrogen atom). Practically the only place in physics where such numbers turn up is in wave phenomena. But the orbiting electrons are also particles. Thus, de Broglie began by trying to make the electron in the atom simultaneously obey both the principle of Fermat and the principle of Maupertuis. The result of this attempt was the de Broglie relation

$$mv = hk \tag{3.1.7}$$

where k is the wave number of the electron wave, h is Planck's constant divided by 2π, m is the mass of the electron (now considered as a particle), and v is the particle's speed (which de Broglie was also able to identify with the group velocity of the waves).[65] Working backwards from the de Broglie relation, it is easy to see that two variational principles (as applied to electrons) are really one and the same. Let us express Fermat's principle in terms of the wave number k (rather than n or the phase velocity v).

$$\delta\int k\,dl = 0. \qquad \text{(optics)} \qquad (3.1.8)$$

Eq. (3.1.8) expresses the condition that, in the geometrical optics limit, small variations in the path do not affect the optical path length. This is why the ray is where it is: only along the actual ray can neighboring bundles of virtual paths interfere constructively. Eq. (3.1.8) must apply, in the geometrical optics limit, to all wave disturbances that obey the superposition principle. To apply eq.(3.1.8), which is very general, to the particular case of matter waves, we use the de Broglie relation, eq. (3.1.7). Eq. (3.1.8) immediately becomes Maupertuis' principle (eq.(3.1.6)). Thus, classical mechanics can be understood as the geometrical-optics limit of wave-mechanics: the particle trajectories are the rays of the matter waves. Maupertuis' principle therefore amounts to a special case of Fermat's principle--the special case involving the matter waves of quantum mechanics. In the context of general relativity, this result may assume deep significance.(See the epilogue at the end).

3.2. On the relation between "F=ma" optics and Hamilton's formulation of the optical-mechanical analogy

We can, if we wish, formulate the problem of the motion of photons and massive particles in the Schwarzschild field in the same way as Hamilton formulated the motion of classical particles and the paths of light in refractive media. In analogy to eq.(3.1.1) or (3.1.3), we define a characteristic function

$$V(y,\ x) \equiv \int_{y}^{x} n^2 v\,dl \qquad (3.2.1)$$

where y and x are two points in space. V will then satisfy the partial differential equation

$$\sum_i \left(\frac{\partial V}{\partial x_i}\right) = n^4 v^2. \qquad (3.2.2)$$

If eq. (3.2.2) can be solved, V will contain all the information we require about the trajectories of photons and particles in the Schwarzschild field. This would lead us into a realm of high abstraction, with no increase in calculating power over eq.(2.2.7).

But it is worth a moment's reflection to see that eq.(3.2.2) is actually equivalent to eq. (2.2.7). We may write

$$dl = \sum_i \cos\theta_i\, dx_i \qquad (3.2.3)$$

where the $\cos\theta_i$ are the direction cosines of dl, i.e., the cosines of the angles dl makes with the Cartesian axes. If we now parameterize the path by a stepping parameter A (yet to be defined), then

$$\cos\theta_i = (dx_i/dA)/(dl/dA). \tag{3.2.4}$$

Then V can be written as

$$V = \sum_i \int n^2 v\left(\frac{dl}{dA}\right)^{-1} \frac{dx_i}{dA} dx_i \tag{3.2.5}$$

from which it follows that

$$\frac{\partial V}{\partial x_i} = n^2 v\left(\frac{dl}{dA}\right)^{-1} \frac{dx_i}{dA}. \tag{3.2.6}$$

Substituting eq.(3.2.5) into eq. (3.2.1), then differentiating both sides of eq.(3.2.1) with respect to A gives

$$\sum_i \frac{d}{dA}\left[n^2 v\left(\frac{dl}{dA}\right)^{-1} \frac{dx_i}{dA}\right]^2 = \sum_i \frac{\partial}{\partial x_i}(n^4 v^2)\ \frac{dx_i}{dA}. \tag{3.2.7}$$

The left-hand side of eq. (3.2.7) cries out for us to define A by

$$\frac{dl}{dA} = n^2 v \tag{3.2.8}$$

just as in eq.(2.2.6). For then eq.(3.2.7) becomes

$$2\sum_i \frac{d^2x_i}{dA^2}\frac{dx_i}{dA} = \sum_i \frac{\partial}{\partial x_i}(n^4v^2)\frac{dx_i}{dA}. \qquad (3.2.9)$$

The direction of the "velocity" dx/dA at a given point can be chosen at will, in the form of initial conditions. Thus, for eq. (3.2.9) to hold in all circumstances, the coefficients of the "velocity" components dx_i/dA must be equal term-by-term:

$$d^2x_i/dA^2 = \partial(1/2\ n^4v^2)/\partial x_i$$

or, in the vector form, eq.(2.2.7) precisely.

Thus, eq. (2.1.6) (and its generalization, eq.(2.2.7)) can be regarded as equivalent to Hamilton's formulation of the optical-mechanical analogy: we need only make the right choice for the stepping parameter. Again, it is to be noted both eq. (3.2.2) and eq. (2.2.7), written in the generalized forms required for motion in the Schwarzschild metric, include classical mechanics (for which n → 1 and A → t) and classical geometrical optics (for which $v \to c_o/n$) as special cases.

Since $dx_i/dA = \partial V/\partial x_i$, the generalized version of Hamilton's differential equation (eq.(3.2.2) becomes, for us, a first integral of the motion (eq.(2.2.8), equivalent to "conservation of energy". Hamilton, of course, would have regarded $dx_i/dA = \partial V/\partial x_i$ as the prescription for extracting answers, in the form of velocity components (dx_i/dA) by differentiation of the solution V, to the differential equation!