Part I

Preliminaries

The Mathematics of Geometrical and Physical Optics: The k-function and its Ramifications. O.N. Stavroudis

ISBN: 3-527-40448-1

1 Fermat's Principle and the Variational Calculus

In the seventeenth century light was believed to be a flow of *corpuscules*, 'little bodies'; their trajectories were called *rays*. Pierre de Fermat asserted that Nature was intrinsically lazy and that those corpuscules 'chose' a trajectory that made their time of transit from point to point a minimum. We refer to this anthropomorphism as *Fermat's Principle*. It was a successful hypothesis. With it, Fermat was able to derive the law of refraction, *Snell's law*, in an economical and precise way.[1]

The connection between Optics and the variational calculus came some years after Fermat when the Swiss mathematician Jacob Bernoulli proposed a problem, the *brachistochrone*, and offered a prize for its solution. Consider a rigid wire connecting a pair of points, fixed in space, on which a bead slides under the force of gravity but without friction. The problem was to find that shape of the wire for which the time of transit of the bead, from one point to the other, was a minimum.[2]

The connection between geometrical optics and Fermat's principle is clear. Jacob's solution was to calculate the vertical force on the bead, taking into account the constraint imposed by the rigid wire. He related this force to an index of refraction function that depended on the height of the bead on the wire. He partitioned the space between the initial and terminal points into horizontal lamina each having a constant refractive index that was determined by its height. Then he could use Snell's law to trace a ray down from the initial point, resulting in a polygonal ray path that approximated the desired solution. As the number of lamina increased and as each thickness approached zero, the polygonal figure approached a continuous curve which was the desired shape of the rigid wire. This curve turned out to be an arc of a cycloid.[3]

Jacob Bernoulli was very pleased with his solution, so much so that he awarded to himself the prize that he had offered, and disregarded the efforts of his brother Jean, who also solved the brachistochrone problem, from an entirely different point of view.

Jean made use of the newly discovered differential calculus and the fact that the first derivative of a function vanishes at its maximum or minimum value. He expressed the time of transit of the bead from the initial point to its terminal point as a an integral of the reciprocal of its velocity. The first derivative of this integral must vanish at a minimum and he obtained conditions that the solution curve must satisfy. Subsequently Leonard Euler extended Jean's

---

[1] Sabra 1967, Chapter V. An account of the history and background of Fermat's principle.

[2] Bliss 1925, pp. 65–72. Caratheodory 1989, pp. 235–236 uses the Hamiltonian which we will encounter in Chapter 3. Woodhouse 1964, Chapters I and II provides a more detailed historical account. Courant & Robbins 1996, pp. 381–384. In Smith 1959, pp. 644–655 there is an English translation of Bernoulli's original paper and announcement.

[3] Bliss 1946, Chapter VI. Jean's use of the calculus in generating the variational calculus.

The Mathematics of Geometrical and Physical Optics: The k-function and its Ramifications. O.N. Stavroudis

ISBN: 3-527-40448-1

method to more general problems and obtained differential equations for their solution. Jean's method can rightfully be called the beginning of modern Calculus of Variations.[4]

It is natural to refer to a solution of a variational problem as an *extremal arc* or more simply as an *extremal*. We will interpret the principle of Fermat in terms of the language of the variational calculus and apply modern mathematics to that basic axiom of geometrical optics and develop it as far as we can.

1.1 Rays in Inhomogeneous Media

We have seen that the basic assumption of geometrical optics is Fermat's principle: A ray path that connects two points in any medium is that path for which the time of transit is an extremum. To be more explicit, out of the totality of all possible paths connecting the two points, A and B, a ray is that unique path for which the time of transit is either a maximum or a minimum. Of course if A and B are conjugates, if B is a perfect image of A, then the ray path is not unique; every ray passing through A must also pass through B.

The time of transit between two points, A and B, is given by the equation

$$T = \int_A^B dt = \int_A^B \frac{ds}{v} = \int_A^B \frac{nds}{c}, \tag{1.1}$$

where c is the velocity of light *in vacuo*, v its velocity in the medium through which it propagates and n the refractive index of that medium. The arc length along the ray or trajectory is s. The optical medium is said to be *homogeneous* if n is constant; it is *inhomogeneous* but *isotropic* if n is a function of position. It is *anisotropic* if the refractive index of the medium depends on the ray's direction.

The convention most used is to drop c from the equations and to use the *optical path length* I, instead of the *time of transit* T, as the variational integral. Thus

$$I = \int_A^B nds. \tag{1.2}$$

In what follows we take the medium to be inhomogeneous so that the refractive index is a function of position $n = n(x,\ y,\ z)$. A possible path connecting A and B is given parametrically by the three coordinate functions $x(t),\ y(t),\ z(t)$ where the choice of the parameter t is entirely arbitrary. If A has the coordinates $(a_1,\ a_2,\ a_3)$ and B, $(b_1,\ b_2, b_3)$ then it must be that

$$\begin{aligned} x(t_0) &= a_1, \quad y(t_0) = a_2, \quad z(t_0) = a_3, \\ x(t_1) &= b_1, \quad y(t_1) = b_2, \quad z(t_1) = b_3, \end{aligned} \tag{1.3}$$

[4] Bliss 1946, Chapter I. Bolza 1961, Chapter 1. Clegg 1968, Chapter 3.

so that

$$I(A,\ B) = \int_{t_0}^{t_1} n(x,\ y,\ z)ds = \int_{t_0}^{t_1} n(x,\ y,\ z)\frac{ds}{dt}dt, \tag{1.4}$$

where the Pythagorean theorem gives us

$$\frac{ds}{dt} = s_t = \sqrt{x_t^2 + y_t^2 + z_t^2}. \tag{1.5}$$

Here, the subscript $(_t)$ denotes differentiation with respect to the parameter t. This subscript notation for both ordinary and partial differentiation will be used extensively in what follows.

In these terms then the problem is to find that curve, given by $x(t),\ y(t),\ z(t)$, for which $I(A,\ B)$ is an extremum.

1.2 The Calculus of Variations

This problem is a special case of a more general problem that belongs to that body of mathematics known as the Calculus of Variations. That more general problem is to find the curve in space, given by $y(x),\ z(x)$ for which the integral

$$I = \int_a^b f\big(x,\ y(x),\ z(x),\ y_x(x),\ z_x(x)\big)dx, \tag{1.6}$$

is an extremum. The function f is always known since it is determined by the nature of the problem; for example, in Eq. 1.4, f is equal to $n(x,\ y,\ z)ds/dt$.

Here we need to find expressions for $y(x)$ and $z(x)$ that make Eq. 1.6 an extremum. First assume that $\overline{y}(x)$ and $\overline{z}(x)$ represent a solution, a curve for which Eq. 1.6 is an extremum. In addition let $\eta(x),\ \zeta(x)$ be any two functions, sufficiently differentiable, such that

$$\begin{aligned} &\eta(a) = \eta(b) = 0, \\ &\zeta(a) = \zeta(b) = 0. \end{aligned} \tag{1.7}$$

Now form a one-parameter family of curves given by

$$y(x) = \overline{y}(x) + h\,\eta(x), \qquad z(x) = \overline{z}(x) + h\,\zeta(x), \tag{1.8}$$

where h is the parameter. By virtue of Eq. 1.7 these curves all pass through the end points of the integral; when the parameter h is *zero* we have, by definition, the solution curve. We replace $y(x)$ and $z(x)$ in the variational integral, Eq. 1.6, by using Eq. 1.8 to get

$$\begin{aligned} I(h) \ &= \int_a^b f\Big(x,\ \overline{y}(x) + h\,\eta(x),\ \overline{z}(x) + h\,\zeta(x), \\ &\qquad \overline{y}_x(x) + h\,\eta_x(x),\ \overline{z}_x(x) + h\,\zeta_x(x)\Big)dx. \end{aligned} \tag{1.9}$$

Because of our construction, if $h = 0$ then I is at an extremum value and, for that value of h, dI/dh must vanish. We calculate this derivative, set it equal to zero and get, from Eq. 1.9

$$\left.\frac{dI}{dh}\right|_{h=0} = \int_a^b \left\{ \frac{\partial f}{\partial y}\eta + \frac{\partial f}{\partial z}\zeta + \frac{\partial f}{\partial y_x}\eta_x + \frac{\partial f}{\partial z_x}\zeta_x \right\} dx = 0. \tag{1.10}$$

Apart from the properties given in Eq. 1.7, the functions $\eta(x)$ and $\zeta(x)$ are entirely arbitrary, a fact that will be important later.

We expand Eq. 1.10 using integration by parts. Recall that,

$$\int_a^b u\, dv = u\, v\Big|_a^b - \int_a^b v\, du,$$

so that

$$\int_a^b \frac{\partial f}{\partial y}\eta\, dx = \left[\eta \int_a^x \frac{\partial f}{\partial y} dx\right]_a^b - \int_a^b \Big[\int_a^x \frac{\partial f}{\partial y} dx\Big]\eta_x dx. \tag{1.11}$$

Since η vanishes at a and b, the first term vanishes. In exactly the same way we get

$$\int_a^b \frac{\partial f}{\partial z}\zeta\, dx = -\int_a^b \Big[\int_a^x \frac{\partial f}{\partial z} dx\Big]\zeta_x dx. \tag{1.12}$$

Substituting Eqs. 1.11 and 1.12 into Eq. 1.10 results in

$$\int_a^b \left\{ \Big[\frac{\partial f}{\partial y_x} - \int_a^x \frac{\partial f}{\partial y} dx\Big]\eta_x + \Big[\frac{\partial f}{\partial z_x} - \int_a^x \frac{\partial f}{\partial z} dx\Big]\zeta_x \right\} dx = 0. \tag{1.13}$$

Note that if the quantities in brackets are constant then the integral vanishes and the condition is satisfied.

This condition is also sufficient. Recall that our choice of the functions η and ζ is completely arbitrary. For the integral to vanish for all possible choices of these functions then the coefficients of their derivatives in Eq. 1.13 must be constant.[5] We conclude that

$$\begin{cases} \dfrac{\partial f}{\partial y_x} - \displaystyle\int_a^x \frac{\partial f}{\partial y} dx = \text{constant} \\ \dfrac{\partial f}{\partial z_x} - \displaystyle\int_a^x \frac{\partial f}{\partial z} dx = \text{constant}. \end{cases} \tag{1.14}$$

[5] Bliss 1946, pp. 10–11 calls this the *Fundamental Lemma of the Calculus of Variations*. I believe that the proof given here is simpler.

If f possesses second derivatives we get the *Euler equations*

$$\begin{cases} \dfrac{d}{dx}\dfrac{\partial f}{\partial y_x} & = \frac{\partial f}{\partial y} \\ \dfrac{d}{dx}\dfrac{\partial f}{\partial z_x} & = \frac{\partial f}{\partial z}, \end{cases} \tag{1.15}$$

a pair of simultaneous *ordinary* differential equations. Recall that f describes the nature of the particular problem and therefore must be known. The solution is an extremal arc that connects the fixed initial and terminal points. Each pair of these end points provide boundary conditions that define a solution. The aggregate of all such solutions to Eq. 1.15 is called a *field of extremals.*

We will need yet another relationship. The total derivative of f with respect to x is

$$\begin{aligned} \frac{df}{dx} &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} y_x + \frac{\partial f}{\partial z} z_x + \frac{\partial f}{\partial y_x} y_{xx} + \frac{\partial f}{\partial z_x} z_{xx} \\ &= \frac{\partial f}{\partial x} + \Big[\frac{d}{dx}\frac{\partial f}{\partial y_x}\Big] y_x + \Big[\frac{d}{dx}\frac{\partial f}{\partial z_x}\Big] z_x + \frac{\partial f}{\partial y_x} y_{xx} + \frac{\partial f}{\partial z_x} z_{xx} \\ &= \frac{\partial f}{\partial x} + \frac{d}{dx}\Big[y_x \frac{\partial f}{\partial y_x} + z_x \frac{\partial f}{\partial z_x}\Big], \end{aligned} \tag{1.16}$$

in which we use the Euler equations, Eq. 1.15 to get

$$\frac{\partial f}{\partial x} = \frac{d}{dx}\Big[f - y_x \frac{\partial f}{\partial y_x} - z_x \frac{\partial f}{\partial z_x}\Big]. \tag{1.17}$$

1.3 The Parametric Representation

The problem can also be expressed in parametric form.[6] We represent the arcs connecting the two end points, A and B, by the coordinate functions $x(t),\ y(t)$ and $z(t)$ of the arbitrary parameter t. It must be that when $t = a$, all possible arcs must pass through A and when $t = b$ they must all pass through B. With this proviso the variational integral becomes

$$I = \int_a^b f\big(x(t),\ y(t),\ z(t),\ x_t(t),\ y_t(t),\ z_t(t)\big) dt. \tag{1.18}$$

It is important, indeed vital, to understand that the parameter t must be applied *uniformly* to all of these possible paths connecting A to B. The *choice* of the parameter t is unimportant and can be anything convenient.

However the choice of t cannot effect the statement of this variational problem and therefore any transformation of t must leave the structure of Eq. 1.18 completely unchanged. To show this [7] we use the *reductio ad absurdum* argument; we assume the contrary and demonstrate a contradiction. First assume that f does indeed depend explicitly on t so that it takes the form

$$f = f\big(t,\ x(t),\ y(t),\ z(t),\ x_t(t),\ y_t(t),\ z_t(t)\big).$$

[6] Bliss 1946, Chapter V. Bolza 1961, Chapter IV. Clegg 1968, Chapter 7.
[7] Bliss 1946 Chapter V. Theorem 41.1.

If we apply a linear transformation to t, say, $t \to \tau + h$, then the variational integrand becomes,

$$f\big(\tau + h, x(\tau + h), y(\tau + h), z(\tau + h), x_\tau(\tau + h), y_\tau(\tau + h), z_\tau(\tau + h)\big)d\tau.$$

Since the differential of τ cannot contain the constant h the transformed variational integrand does not have the same structure as the original version. This contradiction proves that f cannot depend on t explicitly.

We can take this a little further. Suppose the transform involves a factor as in, say, $t \to h\tau$ so that the variational integrand takes the form,

$$f\big(x(h\tau),\ y(h\tau)\ , z(h\tau),\ x_\tau(h\tau)/h,\ y_\tau(h\tau)/h,\ z_\tau(h\tau)/h\big)hd\tau.$$

Compare this expression with the integrand in Eq. 1.18. For this expression to have the same structure as the original variational integrand, f must be a homogeneous function[8] of $x_t,\ y_t,\ z_t$. That is to say,

$$f\big(x,\ y,\ z, \lambda x_t,\ \lambda y_t,\ \lambda z_t\big) = \lambda f\big(x,\ y,\ z, x_t,\ y_t,\ z_t\big).$$

Taking the derivative of this expression with respect to λ, then setting $\lambda = 1$, yields

$$f = x_t\frac{\partial f}{\partial x_t} + y_t\frac{\partial f}{\partial y_t} + z_t\frac{\partial f}{\partial z_t}, \tag{1.19}$$

showing that f must indeed be a homogeneous function in $(x_t,\ y_t,\ z_t)$.

To summarize these results: A variational problem in terms of a parameter t cannot depend on t explicitly; moreover f must be a homogeneous function in $x_t,\ y_t$ and z_t.

In Chapter 5, in which we look at partial differential equations, we will show that a general solution of Eq. 1.19 is obtainable and that the solution is indeed homogeneous; the condition is therefore sufficient as well as necessary. Observe that Eq. 1.19 is the analog of Eq. 1.17 which, in this parametric case, is trivial.

Again we assume a solution, $\overline{x}(t),\ \overline{y}(t),\ \overline{z}(t)$ and choose arbitrary functions $\xi(t), \eta(t),\ \zeta(t)$ that vanish when $t = a$ and when $t = b$, then form the variational integral

$$I(h) = \int_a^b f(\bar{x} + h\xi, \bar{y} + h\eta\ , \bar{z} + h\zeta\ , \bar{x}_t + h\xi_t,\ \bar{y}_t + h\eta_t,\ \bar{z}_t + h\zeta_t)dt. \tag{1.20}$$

We go through the same steps as before and get

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial f}{\partial x_t} = \dfrac{\partial f}{\partial x} \\ \dfrac{d}{dt}\dfrac{\partial f}{\partial y_t} = \dfrac{\partial f}{\partial y} \\ \dfrac{d}{dt}\dfrac{\partial f}{\partial z_t} = \dfrac{\partial f}{\partial z}, \end{cases} \tag{1.21}$$

the Euler equations for the parametric case.

[8] Rektorys 1969, pp. 454–455.

1.4 The Vector Notation

The vector notation simplifies greatly the results obtained for the parametric case. Suppose we have some differentiable function $f(x,\ y,\ z)$. Then its total differential is,

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz.$$

(Of course f can have any number of independent variables but for our purposes *three* is exactly right.) This can be written as a scalar product of two vectors,

$$df = \left(\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial x},\ \frac{\partial f}{\partial x}\right) \cdot (dx,\ dy,\ dz)\,.$$

The left vector we identify as the *gradient* of f

$$\nabla f = \left(\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y},\ \frac{\partial f}{\partial z}\right). \tag{1.22}$$

If we let $\mathbf{V} = (x,\ y,\ z)$ then the total derivative in vector form is

$$df = \nabla f \cdot d\mathbf{V}. \tag{1.23}$$

When cast in vector form the results of the last section assume a much more compact form. We first define the vector function of the parameter t,

$$\mathbf{P}(t) = \big(x(t),\ y(t),\ z(t)\big);$$

its derivative with respect to t must then be,

$$\mathbf{P}_t(t) = \big(x_t(t),\ y_t(t),\ z_t(t)\big),$$

and the variational integral defined in Eq. 1.18 becomes

$$I = \int f(\mathbf{P},\ \mathbf{P}_t)dt. \tag{1.24}$$

Moreover, as was shown in the last section, f must not depend on t explicitly and it must also be homogeneous in $\mathbf{P}_t$.

Next, define two vector gradients according to Eq. 1.22

$$\begin{aligned} \nabla f \ &= \left(\frac{\partial f}{\partial x},\ \frac{\partial f}{\partial y},\ \frac{\partial f}{\partial z}\right), \\ \nabla_t f &= \left(\frac{\partial f}{\partial x_t},\ \frac{\partial f}{\partial y_t},\ \frac{\partial f}{\partial z_t}\right). \end{aligned} \tag{1.25}$$

Applying these to Eq. 1.21 we get the vector form of the Euler equations

$$\frac{d}{dt}\nabla_t f = \nabla f. \tag{1.26}$$

Because f is homogeneous in $\mathbf{P}_t$ it must be that $f = \nabla_t f \cdot \mathbf{P}_t$, this from Eq. 1.19.

In conclusion one might say that the application of the Calculus of Variations consists of two parts; stating the *question* and getting its *answer*. The *question* part is finding the f-function appropriate to the application. The solution to any of the forms of the Euler equations provides the *answer*.

Of course this is only the briefest introduction to the variational calculus. We have discussed here only those elements that are directly relevant to problems that we will encounter subsequently in geometrical optics, such as rays in inhomogeneous media which follows next.

1.5 The Inhomogeneous Optical Medium

Now we apply the version of the Euler equations in Eq. 1.26 to the problem of rays in a medium in which the refractive index is a function of position [9] as indicated in Eqs. 1.4 and 1.5. Evidently $f(\mathbf{P},\,\mathbf{P}_t) = n(\mathbf{P})(ds/dt) = n(\mathbf{P})\sqrt{\mathbf{P}_t^2}$ which establishes f for this particular problem.

We must emphasize that in this context $\mathbf{P}(t)$ is a vector function representing all possible paths in the medium. Our problem is to find those particular paths that satisfy the Euler equations; those are the rays in this medium.

We cannot use s as the parameter in the statement of the variational problem because each possible arc will have a different geometrical length. A requirement for the application of these methods is that the parameter be uniform for all such curves. But s is not uniform so we must use a different parameter, say t, that *is* uniform over all possible arcs. This leads us to the following expressions

$$\nabla f = \sqrt{\mathbf{P}_t^2}\,\nabla n(\mathbf{P}), \qquad \nabla_t f = n(\mathbf{P})\frac{\mathbf{P}_t}{\sqrt{\mathbf{P}_t^2}}. \tag{1.27}$$

Substituting these into Eq. 1.26, the vector form of the Euler equations, we get

$$\frac{d}{dt}\left(n(\mathbf{P})\,\frac{\mathbf{P}_t}{\sqrt{\mathbf{P}_t^2}}\right) = \sqrt{\mathbf{P}_t^2}\,\nabla n(\mathbf{P}). \tag{1.28}$$

But $ds/dt = \sqrt{\mathbf{P}_t^2}$ so that, reverting back to the arc length parameter s, Eq. 1.28 becomes

$$\frac{d}{ds}\left(n\frac{d\mathbf{P}}{ds}\right) = \nabla n. \tag{1.29}$$

This is the *ray equation* for an inhomogeneous medium. Provided that second derivatives exist, it can be expanded further

$$n\mathbf{P}_{ss} + (\nabla n \cdot \mathbf{P}_s)\mathbf{P}_s = \nabla n. \tag{1.30}$$

As always, we use subscripts to signal differentiation.

Equations 1.29 and 1.30 are ordinary differential equations for rays in a medium whose refractive index is a function of position and is continuous and differentiable in the variables $x,\ y,$ and z. An example of such is the *fish eye* of Maxwell which follows.

[9] Stavroudis 1972a, Chapter II. Luneburg 1964, pp. 164–172 discusses the special case where the medium has central symmetry.

1.6 The Maxwell Fish Eye

The ray equation, Eq. 1.29, works very well with Maxwell's fish eye.[10] The eye of a fish operates in water, a medium with a refractive index much higher than that of air, yet its lens is flat. This suggests that the eye of a fish, flat and immersed in a medium with a relatively high refractive index, has a low optical power implying a long back focal distance. Yet the flat structure includes the retina that then requires a short back focal distance. To explain away this paradox Maxwell postulated that the optical medium of the fish eye had a refractive index function in the following form

$$n(\mathbf{P}) = \frac{1}{1+\mathbf{P}^2}, \tag{1.31}$$

so that its gradient is

$$\nabla n = \frac{-2\mathbf{P}}{(1+\mathbf{P}^2)^2}. \tag{1.32}$$

Plugging this into the ray equation, Eq. 1.29, yields

$$\frac{d}{ds}\left[\frac{\mathbf{P}_s}{1+\mathbf{P}^2}\right] = \frac{-2\mathbf{P}}{(1+\mathbf{P}^2)^2}, \tag{1.33}$$

which quickly becomes

$$(1+\mathbf{P}^2)\mathbf{P}_{ss} - 2(\mathbf{P}\cdot\mathbf{P}_s)\mathbf{P}_s + 2\mathbf{P} = 0, \tag{1.34}$$

whose derivative is

$$(1+\mathbf{P}^2)\mathbf{P}_{sss} - 2(\mathbf{P}\cdot\mathbf{P}_{ss})\mathbf{P}_s = 0. \tag{1.35}$$

I do not know whether the fish eye is accurately described by this model or whether fish are even aware of the existence of these equations but as an example of an application of the Calculus of Variations to geometrical optics it will suffice.

We will contemplate these equations further in Chapter 2 which is concerned with the Differential Geometry of Space Curves.

1.7 The Homogeneous Medium

We can use Eq. 1.30 to handle the case where the refractive index n is a constant so that all its derivatives are zero. Then Eq. 1.30 degenerates to

$$\mathbf{P}_{ss} = 0, \tag{1.36}$$

a linear, ordinary differential of order *two* in vector form whose general solution must be

$$\mathbf{P}(s) = \mathbf{A}s + \mathbf{B}, \tag{1.37}$$

where $\mathbf{A}$ and $\mathbf{B}$ are vector constants of integration.

This is clearly a straight line showing us (as if we didn't already know!) that rays in homogeneous, isotropic media are, indeed, the shortest distance between two points.

[10] Luneburg 1964, pp. 172–182. Stavroudis 1972a, Chapter IV.

1.8 Anisotropic Media

In a certain sense the anisotropic medium is an analog of the inhomogeneous medium. In the latter medium the refractive index is a function of position and it can be represented by $n = n(\mathbf{P})$ while in the anisotropic medium it depends on a ray direction[11]. If $\mathbf{P}_s$ is a unit vector in the direction of a ray then it must be that $n = n(\mathbf{P}_s)$, superficially resembling the inhomogeneous medium but making an enormous difference in the variational integral and the Euler equations. Following Eq. 1.24 the variational integrand takes the form

$$f(\mathbf{P}_t) = n(\mathbf{P}_s)ds/dt = n\left(\frac{\mathbf{P}_t}{\sqrt{\mathbf{P}_t^2}}\right)\sqrt{\mathbf{P}_t^2}, \tag{1.38}$$

so that, in Eq. 1.26, $\nabla f = 0$ and the Euler equation becomes

$$\frac{d}{dt}\nabla_t f = \frac{d}{dt}\nabla_t\left[n\left(\frac{\mathbf{P}_t}{\sqrt{\mathbf{P}_t^2}}\right)\sqrt{\mathbf{P}_t^2}\right] = 0. \tag{1.39}$$

The leading component of the gradient is,

$$\begin{aligned}
\frac{\partial f}{\partial x_t} &= \frac{\partial}{\partial x_t}\left[n\left(\frac{\mathbf{P}_t}{\sqrt{\mathbf{P}_t^2}}\right)\sqrt{\mathbf{P}_t^2}\right] \\
&= \sqrt{\mathbf{P}_t^2}\left[\frac{\partial n}{\partial x_s}\frac{\mathbf{P}_t^2 - x_t^2}{(\mathbf{P}_t^2)^{3/2}} - \frac{\partial n}{\partial y_s}\frac{x_t\, y_t}{(\mathbf{P}_t^2)^{3/2}} - \frac{\partial n}{\partial z_s}\frac{x_t\, z_t}{(\mathbf{P}_t^2)^{3/2}}\right] + n\frac{x_t}{\sqrt{x_t^2}} \\
&= \frac{\partial n}{\partial x_s} - \frac{1}{\mathbf{P}^2}x_t\left[x_t\frac{\partial n}{\partial x_s} + y_t\frac{\partial n}{\partial y_s} + z_s\frac{\partial n}{\partial z_s}\right] + n\frac{x_t}{\sqrt{x_t^2}} \\
&= \frac{\partial n}{\partial x_s} - x_s\left[x_s\frac{\partial n}{\partial x_s} + y_s\frac{\partial n}{\partial y_s} + z_s\frac{\partial n}{\partial z_s}\right] + nx_s \\
&= \frac{\partial n}{\partial x_s} - x_s(\mathbf{P}_s\cdot\nabla_s n) + nx_s.
\end{aligned}$$

We do the same thing with the other two partial derivatives in $\nabla_t f$ to get

$$\begin{cases}
\dfrac{\partial f}{\partial x_t} = \dfrac{\partial n}{\partial x_s} - x_s(\mathbf{P}_s\cdot\nabla_s n) + nx_s \\
\dfrac{\partial f}{\partial y_t} = \dfrac{\partial n}{\partial y_s} - y_s(\mathbf{P}_s\cdot\nabla_s n) + ny_s \\
\dfrac{\partial f}{\partial z_t} = \dfrac{\partial n}{\partial z_s} - z_s(\mathbf{P}_s\cdot\nabla_s n) + nz_s,
\end{cases} \tag{1.40}$$

or, in vector form

$$\nabla_t f = \nabla_s n - \mathbf{P}_s(\mathbf{P}_s\cdot\nabla_s n) + \mathbf{P}_s n$$

[11] Avendaño-Alejo and Stavroudis 2002.

$$= \mathbf{P}_s \times (\nabla_s n \times \mathbf{P}_s) + \mathbf{P}_s n. \tag{1.41}$$

From Eq. 1.39 the derivative of Eq. 1.41 must vanish. It follows that there must exist a vector $\mathbf{A}$ that is independent of t (and therefore independent of s) so that

$$\mathbf{P}_s \times (\nabla_s n \times \mathbf{P}_s) + \mathbf{P}_s n = \mathbf{A}. \tag{1.42}$$

The scalar product of this with $\mathbf{P}_s$ yields

$$n = \mathbf{A} \cdot \mathbf{P}_s. \tag{1.43}$$

It follows from this that

$$\nabla_s n = \mathbf{A}. \tag{1.44}$$

This is about as far as we can go without making contact with physical reality; without taking into account the interaction of light with a physical medium.

From Eqs. 1.43 and 1.44 we can get

$$n = \nabla_s n \cdot \mathbf{P}_s, \tag{1.45}$$

a linear, first order partial differential equation that indicates that n must be a homogeneous function. But this is jumping the gun. We will show this and more in Chapter 5 on First Order Partial Differential Equations.

In this chapter we have covered a great deal of territory. We have studied the Calculus of Variations with fixed end points and its parametric representation and then on to a vector notation. This was then applied to inhomogeneous optical media inhomogeneous in general and to Maxwell's fish eye in particular. We have shown that in a homogeneous medium rays are straight lines. In a final brush with anisotropic media in we get inklings of some of the basic flaws in geometrical optics. But we also have laid some foundations on which will be erected new material in subsequent chapters.