INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392 I

SMR/92 - 22

AUTUMN COURSE

ON

VARIATIONAL METHODS IN ANALYSIS AND MATHEMATICAL PHYSICS

20 October - 11 December 1981

SEMINAR

HAMILTON-JACOBI THEORY IN THE CALCULUS OF VARIATIONS

R.B. MISRA
Department of Mathematics
University of Allahabad
Allahabad
India

HAMILTON-JACOBI THEORY IN THE CALCULUS OF VARIATIONS

by

R. B. MISRA,

Department of Mathematics, University of Allahabad, Allahabad (INDIA).

§ 0. Introduction:

Although the calculus of variations is, in the first instance, concerned with curves (or subspaces) which afford extreme values to given single (or multiple) integrals, it would seem that this unifying effect is not so much a direct consequence of such 'extremal' aspects but is due to the properties of certain congruences of curves (or families of subspaces) associated with variational problems. The study of these congruences and families is the subject matter of the Hamilton-Jacobi theory.

The problems involving single integral are the simplest variational problems in the theory. There are two different aspects of the theory resulting from such integrals: (i) which are not invariant under transformations of the parameter (the so-called non-homogeneous case), and (ii) those which are based on parameter-invariant integrals (the homogeneous case). This distinction is reflected in the fundamental difference between non-relativistic mechanics and quantum mechanics on the one hand, and relativistic particle dynamics and quantum mechanics on the other. This mode of classification appears to be useful from the point of view of physical applications while at the same time direct access is gained to generalized metric differential geometry such as the theory of Finsler spaces.

Similar invariance properties of multiple integrals lead to the differential geometry of areal spaces, Cartan spaces, the rigorous formulation of the theorem of Noether, and to the modern physical field theories.

§ 1. Formulation of single integral problem:

Let t, x^i, $(i = 1, 2, \dots, n)$, be $n+1$ independent real variables representing the coordinates of an $(n+1)$-dimensional real space R_{n+1}. Let P_1 and P_2, with respective coordinates (t_1, x_1^i) and (t_2, x_2^i), be two distinct points of R_{n+1} such that $t_1 < t_2$. We consider a set of n continuous functions

(1.1) $$x^i = x^i(t), \qquad (i = 1, 2, \dots, n),$$

of class, at least, C^1 and satisfying the conditions

(1.2) $$x^i(t_1) = x^i_1, \quad x^i(t_2) = x^i_2.$$

Thus, the functions so defined give parametric representation of a 1-dimensional locus in R_{n+1}, namely a curve say C, joining the points P_1 and P_2. Having assumed the differentiability of the functions x^i we may also construct their first order derivatives with respect to the parameter t:

(1.3) $$\dot{x}^i \overset{\text{def}}{=} dx^i/dt,$$

which are also functions of t.

The $(2n+1)$ variables $(t, x^i, \dot{x}^i)$ represent a line-element of C and, indeed, are regarded as coordinates in $(2n+1)$-dimensional space R_{2n+1}. Now we assume a function $L(t, x^i, \dot{x}^i)$ of these $(2n+1)$ arguments defined at all points of a simply-connected region say G of R_{2n+1}. Hereafter, our all considerations are restricted to this region only. We also assume that the function $L(t,x^i,\dot{x}^i)$ possesses continuous second order derivatives with respect to all its $2n+1$ arguments. The integral of function L along C from P_1 to P_2:

$$(1.4) \qquad I = \int_{P_1}^{P_2} L(t,x^i,\dot{x}^i)\,dt$$

is well-defined and is dependent on the choice of the curve C.

The simplest problem in the calculus of variations is to account for those curves for which the integral (1.4) assumes extreme values as compared with neighbouring curves. And as such one also endeavoures to determine the conditions which the functions x^i's (defining these curves) must satisfy in order that they afford such extreme values to (1.4).

Definition 1.1. The function $L(t,x^i,\dot{x}^i)$ so assumed is called the Lagrangian or fundamental function of the Hamilton-Jacobi theory in the calculus of variations; and the integral (1.4) is called the fundamental integral.

Agreement 1.1. It will always be assumed that the Lagrangian is invariant under arbitrary transformations of the coordinates x^i whose Jacobian does not vanish.

§2. Generalization of above problem:

2.1. Variable end-points: The simplest problem in the calculus of variations as stated in §1 depends, in its formulation, on the fixed end-points P_1 and P_2. A natural extension of the problem is obtained by the requirements that one (or both) of these end-points be variable on a suitable sub-manifold of space R_{n+1} of the variables (t,x^i).

2.2. Multiple integral problems: The single parameter t

in the fundamental integral (1.4) may be replaced by a set of, say, m independent variables $t^1, t^2, \dots, t^m$ (where it is already assumed that the n independent functions x^i depend on these m arguments t^α, $(\alpha = 1,2,\dots,m)$. The problems resulting from this consideration represent a very natural and extremely important generalization of the problem defined in §1. These problems are of particular interest to the field theories of modern physics and in the geometry of areal spaces.

2.3. Lagrangians involving higher order derivatives: In §1 we have observed that the Lagrangian depends solely on the first order derivatives of the functions x^i's which define the curve C. An obvious generalization is suggested by taking the Lagrangians which involve higher order derivatives of the functions x^i's as well. For instance, a Lagrangian $L(t,x^i,\dot{x}^i,\ddot{x}^i)$ involving the derivatives of x^i's upto second order gives rise to the geometry of Kawaguchi spaces.

2.4. The problem of Lagrange: Returning back to the fundamental integral (1.4) the foremost problem is to look for certain subsidiary conditions satisfied by curves so that this integral may be of extreme values. These conditions are generally expressed in terms of equations involving the variables t, x^i and $\dot{x}^i$ which need not be integrable. The resulting problem in the calculus of variations is called a problem of Lagrange.

2.5. More general assumptions regarding the continuity and differentiability properties of both the Lagrangian and the admissible curves (or surfaces).

2.6. The calculus of variations in the large: The classical calculus of variations is primarily concerned with the immediate neighbourhood of a minimizing arc. But, in the so-called 'Calculus of variations in the large', which represents one of the most important modern development of the subject, one is concerned with the complete manifold on which the variational problem is given.

2.7. Direct methods in the calculus of variations: The application of variational methods of the solution of eigenvalue problems represents one of the most important features of the modern analysis.

2.8. Variational methods in the problems of optimal control and numerical methods.

§3. Non-homogeneous cases of simplest variational problem:

Turning back to the simplest variational problem as stated in the §1 we observe that when the Lagrangian $L(t, x^i, \dot{x}^i)$, defining the fundamental integral (1.4), also depends on the parameter t it does not possess certain homogeneous properties. Such non-homogeneous cases are dealt here with special reference to geometrical optics and non-relativistic mechanics.

3.1. Geometrical optics: We consider a 3-dimensional Euclidean space of classical physics equipped with a coordinate system (t, x^i), $(i = 1, 2)$, where t is merely a spatial coordinate. Let $P(t, x^i)$ be a point in an optical medium. A direction through the point P may be specified by $\dot{x}^i \equiv dx^i/dt$. The velocity of a ray of light in that direction will, in general, depend upon the five variables $(t, x^i, \dot{x}^i)$ and is, therefore, denoted by $v(t, x^i, \dot{x}^i)$. If c denotes the velocity of light the refractive index is defined by

$$N(t, x^i, \dot{x}^i) \overset{\text{def}}{=} c\, v^{-1}(t, x^i, \dot{x}^i). \tag{3.1}$$

Definition 3.1. The medium is called isotropic (respectively homogeneous) if the refractive index N happens to be independent of the directional arguments $\dot{x}^i$ (respectively positional coordinates (t, x^i)).

Let $C : x^i = x^i(t)$ represent the path of a ray of light between two points P_1, P_2 corresponding to the parametric values $t = t_1$ and $t = t_2$ respectively. The time T taken to describe the segment P_1P_2 of C is given by

$$T = \int_{t_1}^{t_2} v^{-1}\, ds = \int_{t_1}^{t_2} v^{-1} \{(dx^1)^2 + (dx^2)^2 + (dx^3)^2\}^{1/2} \tag{3.2}$$

in rectangular coordinates. (It may be noted that for the third coordinate t of the space so chosen we have written x^3 here. Also, ds represents the elementary arc-length P_1P_2 of C.) Using (3.1) and replacing x^3 by t above equation reduces to

$$T = \int_{t_1}^{t_2} L(t, x^i, \dot{x}^i)\, dt,$$

where we have put

$$L \overset{\text{def}}{=} c^{-1} N \{1 + (\dot{x}^1)^2 + (\dot{x}^2)^2\}^{1/2}. \tag{3.3}$$

Definition 3.2. A function $H(t, x^i, \dot{x}^i)$ defined by

$$H(t, x^i, p_i) \overset{\text{def}}{=} -L(t, x^i, \dot{x}^i) + \dot{x}^i p_i \tag{3.4}$$

is called the Hamiltonian function, where $p_i \equiv \partial L/\partial \dot{x}^i$.

Theorem 3.1. In order that a family of hypersurfaces

$$S(t, x^i) = \text{Const.} \tag{3.5}$$

be geodesically equidistant it is necessary and sufficient that the function S must satisfy the equation

$$\frac{\partial S}{\partial t} + H(t, x^i, \partial S/\partial x^i) = 0, \tag{3.6}$$

which is called the Hamilton-Jacobi equation.

With a Lagrangian of the type (3.3) interpreted in terms of the refractive index of an optical medium which need be neither isotropic nor homogeneous there also holds the:

Theorem 3.2. Each solution of the Hamilton-Jacobi equation represents a family or succession of wave-fronts, and the canonical congruence belonging to this family represents the corresponding light rays.

Remark 3.1. In optical theory the Hamilton-Jacobi equation (3.6) is frequently referred as the Eikonal equation.

3.2. Non-Relativistic mechanics: The simplest problem of calculus of variations, as stated in §1, is revisited here under the assumption that the Lagrangian depends upon the parameter t which is now the time parameter. How the Hamilton-Jacobi theory points the way to quantum mechanics in the transition to the latter from classical mechanics is the subject matter of this sub-section. This transition is carried through two different ways giving rise to Schrödinger and Heisenberg's schemes. However, at a later stage, the equivalence of two schemes is also established.

3.2.1. Schrödinger scheme: As the range of application of 'geometrical optics' is strictly limited while the theory possessing the phenomena of physical optics such as interference or diffraction has to be based on a second-order partial differential equation which can be expressed physically or geometrically in terms of the undulatory theory of light. Thus, the theory of geometrical optics based on a first-order differential equation, i.e. the Hamilton-Jacobi or Eikonal equation is to be

replaced by a theory based on a second-order partial differential equation, i.e. a wave equation. Moreover, from a purely mathematical point of view the disciplines of geometrical optics and classical mechanics are essentially equivalent at least until the physical interpretation of some relevant mathematical entities is sought. It is natural, therefore, to see whether this transition from classical mechanics to quantum (or non-relativistic) mechanics is possible in a manner directly analogous to the transition from geometrical to physical optics. And, indeed, it is possible if we regard the families of geometrically equidistant hypersurfaces of Hamilton-Jacobi theory as wave-fronts. In analogy with the physical optics there emerges a second-order wave equation called the Schrödinger equation and the undulatory concepts accompanying this equation form the basis of non-relativistic wave mechanics in the so-called Schrödinger representation. Furthermore, the physical quantities in this scheme, which are just ordinary functions of real variables in the classical theory, are now replaced by operators, for instance, differential operators, or more generally operators acting on the elements of certain function spaces such as Hilbert spaces.

3.2.2. Heisenberg's scheme: Beginning with the theory of canonical transformations of classical mechanics with particular reference to Poisson bracket relations it is observed that there are certain properties of these relations which are directly analogous to some of the identities satisfied by the commutators of the quantum mechanical operators of the Schrödinger scheme. This analogue gives rise to Heisenberg's scheme of non-relativistic quantum mechanics.

§ 4. Homogeneous cases of simplest variational problem.

In the preceding section we were concerned with variational problems involving Lagrangians of the type $L(t, x^i, \dot{x}^i)$ which also depend upon the parameter t. But, in metric differential geometry we observe that not only the length of any curve is invariant on the choice of the parameter used in its parametric representation rather one may also deal with the variational problem involving closed curves. Thus, the need for a theory of parameter-invariant problems is inevitable. To summarize, the theory of simplest variational problems involving a Lagrangian independent of the parameter t, used therein, paves the way to metric differential geometries such as that of Finsler (including Riemannian and Minkowskian geometries as special cases).

In the following we demonstrate how the parameter-invariance of the Lagrangian implies its homogeneous character in its directional arguments $\dot{x}^i$'s. Instead of considering a configuration space R_{n+1} (as taken in §1) we begin with an n-dimensional manifold X_n of n variables x^i, $(i = 1, 2, \ldots, n)$. As before, a set of n equations

(4.1) $$x^i = x^i(\tau),$$

where τ is an arbitrary parameter, represents a curve C in X_n. Taking the functions x^i's of class, at least C^1, we form their derivatives

$$\dot{x}^i \equiv dx^i/d\tau$$

determining the components of the tangent vector to C. Now we consider a function $L(x^i, \dot{x}^i)$ of class, at least, C^2 in its all $2n$ arguments. Analogous to (1.4), the integral

(4.2) $$I = \int_{\tau_1}^{\tau_2} L(x^i, \dot{x}^i)\, d\tau,$$

may represent the arc-length of the curve C from the point P_1 to P_2 which correspond to C for the parametric values τ_1 and τ_2.

The problem of calculus of variations resulting from this integral can be formulated as before. The only additional condition to which the present Lagrangian should satisfy is its invariance on the parameter τ. Let

(4.3) $$\sigma = \sigma(\tau)$$

be an arbitrary transformation of the parameter τ, where the function σ is of class C^1 and satisfies

(4.4) $$\dot{\sigma} \equiv d\sigma/d\tau > 0.$$

Under the parameter transformation (4.3) the integral (4.2) takes the form

$$I = \int_{\sigma_1}^{\sigma_2} L\left(x^i, \frac{dx^i}{d\sigma}\dot{\sigma}\right)\frac{d\sigma}{\dot{\sigma}},$$

which, by hypothesis, must be identical with the integral

$$\int_{\sigma_1}^{\sigma_2} L\left(x^i, \frac{dx^i}{d\sigma}\right) d\sigma$$

for any curve C. Clearly, this is possible only if there holds

(4.5) $$L(x^i, \lambda \dot{x}^i) = \lambda L(x^i, \dot{x}^i)$$

for any positive number λ. The last relation characterizes the positively homogeneous character of the Lagrangian in the coordinates $\dot{x}^i$'s. Conversely, when the condition (4.5) holds good the Lagrangian becomes invariant on the choice of the parameter.

Apart from this distinguishing character of the present theory as compared to the one discussed in the preceding sections one also observes that the Hamiltonian function, as defined by equation (3.4), vanishes in the present (homogeneous) case of the Lagrangian. Thus, to get a suitable Hamiltonian function in the present theory it is desirable to seek some suitable canonical variables. For this, we construct the so-called fundamental (or metric) tensor with the following components:

$$g_{ij}(x^k, \dot{x}^k) \overset{\text{def}}{=} \tfrac{1}{2} \dot{\partial}_i \dot{\partial}_j L^2(x^k, \dot{x}^k), \tag{4.6}$$

where we have used the notation $\dot{\partial}_i \equiv \partial/\partial \dot{x}^i$. Transvecting this equation by $\dot{x}^i \dot{x}^j$ and using the homogeneous properties of the Lagrangian we derive the identity

$$g_{ij}(x^k, \dot{x}^k)\, \dot{x}^i \dot{x}^j = L^2(x^k, \dot{x}^k). \tag{4.7}$$

This suggests an alternate form of the integral (4.2):

$$I = \int_{\tau_1}^{\tau_2} \sqrt{\{g_{ij}(x^k, \dot{x}^k)\, \dot{x}^i \dot{x}^j\}}\, d\tau, \tag{4.8}$$

implying

$$(ds)^2 = g_{ij}(x^k, \dot{x}^k)\, dx^i dx^j, \tag{4.9}$$

where ds denotes the arc-length of the curve C measured from P_1 to P_2.

Definition 4.1. An n-dimensional space X_n endowed with a metric (4.9) is called a Finsler space.

Particular cases: The space X_n is called Riemannian (respectively Minkowskian) if the metric tensor components g_{ij} are independent of the directional coordinates $\dot{x}^i$'s (respectively the positional coordinates x^i's).

Note 4.1. It should be noted that the present terminology of Minkowskian space has nothing to do with the 4-dimensional Minkowski space of the special theory of relativity.

§5. References:

[1] Bliss, G.A.: Lectures on the Calculus of Variations. Univ. of Chicago Press, Chicago (1946).

[2] Busemann, H.: The Geometry of Geodesics. Academic Press, New York (1955).

[3] Carathéodory, C.: Variationsrechnung und partielle Differentialgleichungen erster Ordnung. Teubner, Leipzig-Berlin (1935).

[4] Dirac, P.M.: Principles of Quantum Mechanics. Clarendon Press, Oxford (1958).

[5] Hamilton, W.R.: Mathematical Papers, Vol. I. Geometrical Optics, Cambridge Univ. Press, Cambridge (1931).

[6] Herzberger, M.: Strahlenoptik. Springer, Berlin (1931).

[7] Hund, F.: Materie als Feld. Springer, Berlin-Göttingen-Heidelberg (1954).

[8] Kawaguchi, A.: Die Differentialgeometrie in der verallgemeinerten Mannigfaltigkeit. Rend. Circ. Mat. Palermo 56 (1932), 245-276.

[9] ——: On areal spaces I, II, III. Tensor (N.S.) 1 (1951), 14-45, 67-88, 89-103.

[10] Lanczos, C.: The Variational Principles of Mechanics. Univ. of Toronto Press, Toronto (1949).

[11] Morse, M.: The Calculus of Variations in the Large. Amer. Math. Soc. Colloq. Publ., Vol. 18, A.M.S, New York (1934).

[12] Polak, L.S.: Variational Principles of Mechanics (Russian). Gosudarstv. Izdat. Fiz.-Mat. Lit., Moscow (1959).

[13] Prange, G.: Die allgemeinen Integrationsmethoden der analytischen Mechanik. Enzykl. math. Wiss. 4, Teubner, Leipzig (1935).

[14] Rund, H.: The Differential Geometry of Finsler Spaces. Springer, Berlin - Göttingen - Heidelberg (1959).

[15] ——: Note on the Lagrangian formalism in relativistic mechanics. Nuovo Cimento 23 (1962), 227-232.

[16] —: Canonical formalism and the relativistic wave equation. Perspectives in Geometry and Relativity, Univ. of Indiana Press (1966).

[17] —: Variational Principles for combined vector and tensor fields. Abh. math. Sem. Univ. Hamburg 29 (1966), 243-262.

[18] ——: The Hamilton-Jacobi Theory in the Calculus of Variations. Van Nostrand, London - New York - Toronto (1966).

[19] Synge, J.L.: Geometrical Mechanics and de Broglie Waves. Cambridge Univ. Press, Cambridge (1954).

[20] ——: Relativity: the Special Theory. North-Holland, Amsterdam (1955).

[21] Teach, V. B.: The Hamilton-Jacobi Theory for the Problems of Lagrange in Parametric Form. Contributions to the Calculus of Variations, 1933-37, 165-206, Univ. of Chicago Press, Chicago (1938).

[22] Tonelli, L.: Fondamenti di calcolo delle variazioni, Vol. 1, 2, Zanichelli, Bologna (1921, 1923).

[23] Vanstone, J.R.: The Hamilton-Jacobi equations for a relativistic charged particle. Canadian Math. Bull. 6 (1963), 341-349.

———x———