CANONICAL THEORIES OF LAGRANGIAN DYNAMICAL SYSTEMS IN PHYSICS

H.A. KASTRUP

Institut für Theoretische Physik, RWTH Aachen, 51 Aachen, FR Germany

NORTH-HOLLAND PHYSICS PUBLISHING–AMSTERDAM

PHYSICS REPORTS (Review Section of Physics Letters) 101, Nos. 1 & 2 (1983) 1–167. North-Holland, Amsterdam

CANONICAL THEORIES OF LAGRANGIAN DYNAMICAL SYSTEMS IN PHYSICS

H.A. KASTRUP

Institut für Theoretische Physik, RWTH Aachen, 51 Aachen, FR Germany

Received May 1983

Contents:

Single orders for this issue

PHYSICS REPORTS (Review Section of Physics Letters) 101, Nos. 1 & 2 (1983) 1–167.

Copies of this issue may be obtained at the price given below. All orders should be sent directly to the Publisher. Orders must be accompanied by check.

Single issue price Dfl. 93.00, postage included.

0 370-1573/83/0000–0000/$50.10

Abstract:

By reformulating the variational problem for a given classical Lagrangian field theory in the framework of differential forms, one can show (Lepage) that for $m \geq 2$ independent and for $n \geq 2$ dependent (field) variables $z^a = f^a(x)$ a much wider variety of Legendre transformations $v^a_\mu = \partial_\mu f^a(x) \rightarrow p^\mu_a$, $L \rightarrow H$, exists than has been employed in physics. The different canonical theories for a given Lagrangian can be classified according to the rank of the corresponding basic canonical m-form.

Each such canonical theory leads to a Hamilton–Jacobi theory, the "wave fronts" of which are transversal to solutions of the field equations.

Two canonical theories are discussed in more detail: The one by DeDonder and Weyl which employs the conventional canonical momenta $p^\mu_a = \partial L/\partial v^a_\mu$ and the more sophisticated one by Carathéodory, the HJ theory of which is more intimately related to that of mechanics than the conventional one.

Generalizing results from mechanics one can show that each solution of a HJ equation which depends on a parameter generates a conserved current for those extremals which are transversal to that wave front.

The geometrically very rich, but algebraically rather complicated canonical formalism of Carathéodory provides interesting new approaches for the "qualitative" analysis of classical field theories. For instance: solutions of the field equations which give a vanishing Lagrangian density L are associated with singularities in the transversality relations between wave fronts and extremals.

A number of examples (strings, gauge theories etc.) illustrates the wealth of possible physical applications of these more general canonical formalisms for field theories, which, up to now, have been ignored almost completely by physicists.

Introduction

Some years ago, when I tried to understand the papers by Dashen, Hasslacher and Neveu [1974, I, II] on semiclassical approximations in quantum field theories, I started wondering, why Hamilton–Jacobi (=HJ) theories for classical fields were never used in physics and whether they did exist at all. Looking up the mathematical literature I discovered:

Not only did a wealth of papers on HJ theories for fields exist, but in addition there were important results concerning the canonical formulation of classical field theories which have been completely ignored by the general physics community! In fact, whereas the mathematical theory of the calculus of variations for systems with *one* independent variable has had a very strong influence on the development of mechanics, quantum mechanics and quantum field theory, the mathematical papers on the calculus of variations for systems with *several* independent variables have left almost no traces within the modern developments of field theories in physics.

One of the main reasons for this development is, of course, that physicists to a large extent consider field theories as (quantum) mechanical systems with an infinite number of degrees of freedom. This approach, which has been extremely fruitful and which is, of course, completely justified from a very appealing point of view of physics and functional analysis, for a long time had the tendency to ignore the rich geometrical structure of classical solutions of partial differential equations which serve as the starting point for a "corresponding" quantum field theory (this remark does not apply to General Relativity). However, during the last years, in view of the successes of gauge theories in particle physics, we have learnt again to appreciate the geometrical aspects of field theories.

I am convinced that essential parts of those geometrical properties still have to be discovered and applied by physicists and that the large variety of canonical theories with their associated HJ equations for a given Lagrangian field theory does represent such an undiscovered part and it is the aim of this article to draw attention to these aspects and to illustrate them by physical examples.*

Volterra [1890] was the first to generalize the concepts which Hamilton and Jacobi had developed for optics and mechanics to a field theory with 2 independent variables and to write down a "HJ" equation for such a system. The subject was taken up by Fréchet [1905] who treated the case of m independent and n dependent variables and who, in addition, generalized another important property of a solution $S(t, q)$, $q = (q^1, \ldots, q^n)$ of a HJ equation in mechanics [see, e.g. Whittaker, 1959, p. 324]: Let $S(t, q; a)$ be a solution which depends on a parameter a. Then the quantity $G = \partial S/\partial a$ is a constant of motion "along" an extremal with canonical coordinates $(q^1(t), \ldots, q^n(t), p_1(t), \ldots, p_n(t))$, for which the relations $p_j(t) = \partial_j S(t, q(t); a)$, $\partial_j := \partial/\partial q^j$, hold. These HJ constants of motion constitute a larger class of conserved quantities than those obtained from Noether's theorem [Noether, 1918], which is a special case of Jacobi's one (for details see section 4.2)! Whereas Noether's theorem provides conserved quantities for *any* solution of the equations of motion, Jacobi's theorem asserts the possibility of additional constants of motion for *special* solutions which depend on certain parameters not associated with a general invariance group of the Lagrangian.

In the case of field theories a solution of the HJ equation depending on a parameter provides a conserved current $G^\mu(x)$, $\mu = 1, \ldots, m$, $\partial_\mu G^\mu = 0$ associated with a solution of the field equations which is "transversal" to that HJ "wave front" (more details below).

The work of Volterra and Fréchet was summarized in 1911 by the Belgian mathematician DeDonder [1911].

In his famous address during the 2nd international congress of mathematicians in 1900 in Paris Hilbert had discussed the importance of a certain path-independent integral – which now bears his name – for the calculus of variations with one independent variable [Hilbert, 1900, 1906]. A few years later Mayer [1904, 1906] recognized and analyzed the important relationship of Hilbert's independent integral to the theory of Hamilton and Jacobi. In the wake of this development DeDonder in 1913 introduced an independent integral and an associated "HJ" equation for variational problems with several independent variables. In 1930 DeDonder published a monograph on the subject. This theory was discussed and analyzed further by Weyl in 1934/35 and since then bears the name of DeDonder and Weyl (DW). Its essential features are as follows:

Let $L(x, z, v)$ be the Lagrangian (density) of a system, where $x = (x^1, \ldots, x^m)$ are the independent variables, $z = (z^1, \ldots, z^n)$ the n variables which become dependent variables (functions) $z^a = f^a(x)$ on the extremals and $v = (v^1_1, \ldots, v^a_\mu, \ldots, v^n_m)$ the variables which become $v^a_\mu = \partial_\mu f^a(x)$, $\mu = 1, \ldots, m$, $a = 1, \ldots, n$ on the extremals. The canonical momenta in the DW theory are defined by $\pi^\mu_a := \partial L/\partial v^a_\mu$ and the invariant (!) Hamilton function is the Legendre transform $H = \pi^\mu_a v^a_\mu - L$. The DWHJ equation is the 1st order partial differential equation

$$\partial_\mu S^\mu(x, z) + H(x, z, \pi^\mu_a = \partial_a S^\mu) = 0\,, \tag{I,1}$$

$$\pi^\mu_a = \partial_a S^\mu(x, z) := \partial S^\mu/\partial z^a\,, \tag{I,2}$$

for the m functions $S^\mu(x, z)$, $\mu = 1, \ldots, m$.

If $S^\mu(x, z; a)$, $\mu = 1, \ldots, m$, are solutions of eq. (I,1) – in the following we shall speak of the

* I call a dynamical system a Lagrangian one, if its evolution equations can be derived from a Lagrangian function (density).

"solution $S^\mu(x, z)$" – and $z^b = f^b(x)$, $b = 1, \ldots, n$, solutions of the Euler–Lagrange equations for which the relations $\pi^\mu_b(x) = \partial_b S^\mu(x, z = f(x); a)$ hold, then the components $G^\mu(x)$ of the conserved current mentioned above are given by $G^\mu(x) = (\partial S^\mu/\partial a)\,(x, z = f(x))$.

In mechanics one has the important notion of a "complete" integral: Suppose there exists a solution $S(t, q; a)$ of the HJ equation which depends on n constants a_j, $j = 1, \ldots, n$, such that $\det(\partial^2 S/\partial q^j \partial a_k) \neq 0$, then the solutions $q^j(t) = f^j(t; a, b)$ of the n equations $(\partial S/\partial a_j)(t, q, a) = b^j = \text{const.}$ are extremals, with $p_j(t) = \partial_j S(t, q = f(t; a, b), a)$. As these functions depend on $2n$ arbitrary parameters, they constitute the most general solution of the equations of motion.

Similarly, suppose a solution $S^\mu(x, z; a)$ of eq. (I,1) depends on mn parameters a^ν_c, $\nu = 1, \ldots, m$, $c = 1, \ldots, n$, then $S^\mu(x, z; a)$ is called a "complete" integral, if $\det(\partial^2 S^\mu/\partial z^b \partial a^\nu_c) \neq 0$. It is then again possible to construct solutions of the Euler–Lagrange equations if certain integrability conditions – which do not exist in mechanics – are satisfied: If $S^\mu(x, z)$ is a solution of the DWHJ eq. (I,1), then we have, according to eqs. (I,2), $\pi^\mu_a = \partial_a S^\mu(x, z) =: \psi^\mu_a(x, z)$. Performing the Legrendre transformation $\pi^\mu_a \to v^a_\mu$, we obtain "slope" functions $v^a_\mu = \phi^a_\mu(x, z)$ which can only be identified with derivatives $\partial_\mu f^a(x)$ of functions $z^a = f^a(x)$, if the integrability conditions

$$\frac{\mathrm{d}}{\mathrm{d}x^\nu} v^a_\mu := \partial_\nu \varphi^a_\mu + \partial_b \varphi^a_\mu \cdot \varphi^b_\nu = \frac{\mathrm{d}}{\mathrm{d}x^\mu} v^a_\nu = \partial_\mu \varphi^a_\nu + \partial_b \varphi^a_\nu \cdot \varphi^b_\mu$$

are fulfilled. These conditions impose severe restrictions on the solutions $S^\mu(x, z)$ which in general are harder to solve than the DWHJ eq. (I,1) itself.

There is another problem associated with the DW "wave fronts" $S^\mu(x, z)$ which does not exist in mechanics: For a mechanical system with n degrees of freedom the wave fronts, transversal to a family of extremals, are given by $S(t, q) = \sigma = \text{const.}$ and are therefore n-dimensional in general. This is no longer the case for the DWHJ theory where the transversal wave fronts for $n, m \geq 2$ are given by the equations $S^\mu(x, z) = \sigma^\mu = \text{const.}$, $x^\mu = \text{const.}$, $\mu = 1, \ldots, m$, that is to say, the DWHJ wave fronts in general are $(n - m)$-dimensional.

This last "defect" does not exist in the HJ theory for fields invented by Carathéodory in 1929. In this theory the wave fronts transversal to the extremals are n-dimensional as in mechanics. However, Carathéodory's "Legendre"-transformation $v^a_\mu \to p^\mu_a$, $L \to H_c$ is more complicated:

$$p^\mu_a = (-L)^{1-m} \bar{T}^\mu_\rho \pi^\rho_a\,, \qquad H_c = (-L)^{1-m} |T|\,, \tag{I,3}$$

where

$$T^\mu_\nu = \pi^\mu_a v^a_\nu - \delta^\mu_\nu L\,, \qquad \bar{T}^\mu_\rho T^\rho_\nu = \delta^\mu_\nu |T|\,, \qquad |T| := \det(T^\mu_\nu)\,.$$

The associated HJ equation is

$$|(\partial_\nu S^\mu)| + H_c(x, z, p) = 0\,, \tag{I,4}$$

$$\partial_\rho S^\mu p^\rho_a - |(\partial_\nu S^\rho)|\, \partial_a S^\mu = 0\,. \tag{I,5}$$

Because of its highly nonlinear structure Carathéodory's canonical theory for fields does not have much appeal as regards calculational simplicity! However, it has a number of very intriguing structural properties which deserve attention:

(i) For $n \geq 2$ it is the only canonical theory for fields which allows for the same transversality structure of extremals and wave fronts as one encounters in mechanics.

(ii) Its Hamilton function H_c is essentially the determinant of the canonical energy-momentum tensor (T^{μ}_{ν}), a property which is intuitively very appealing. Notice, that for $m = 1$ the expressions (I,3–5) reduce to those in mechanics!

(iii) Given $m - 1$ functions $S^2(x, z), \ldots, S^m(x, z)$ which obey the "transversality" conditions (I,5), the canonical transformation $x^1 \to \hat{x}^1 = x^1$, $x^{\bar{\mu}} \to \hat{x}^{\bar{\mu}} = S^{\bar{\mu}}(x, z)$, $\bar{\mu} = 2, \ldots, m$, $z^a \to \hat{z}^a = z^a$, has the following properties: In the new frame the Hamilton function $\hat{H}_c$ is given by $\hat{H}_c = \hat{T}^1_1$ and the canonical field equations take the "mechanical" form

$$\frac{d\hat{p}^1_a}{d\hat{x}^1} = -\frac{\partial \hat{H}_c}{\partial \hat{z}^a}, \qquad \frac{d\hat{z}^a}{d\hat{x}^1} = \frac{\partial \hat{H}_c}{\partial \hat{p}^a_1}, \qquad \hat{p}^{\bar{\mu}}_a = 0, \qquad \bar{\mu} = 2, \ldots, m,$$

$$\partial_1 \hat{S}^1 + \hat{H}_c = 0, \qquad \hat{p}^1_a = \partial_a \hat{S}^1,$$

i.e. on the surfaces $S^{\bar{\mu}}(x, z) = \text{const.}$, $\bar{\mu} = 2, \ldots, m$, the dynamical "flow" of the fields reduces to a "mechanical" one. I consider this property of Carathéodory's canonical theory to be of great importance. It was discovered by E. Hölder [1939].

(iv) At a point $(x, z) \in \mathrm{R}^{m+n}$ the tangent space of an extremal in general will be spanned by m linearly independent tangent vectors and the tangent space of a transversal wave front by n linearly independent tangent vectors. The necessary and sufficient condition for these $m + n$ tangent vectors to be linearly independent is $H_c L \neq 0$. Thus, for solutions of the field equations which have $L = 0$ the transversality properties of extremals and wave fronts become singular (caustics!). The first order condition $L = 0$ can have quite surprising physical implications [Kastrup, 1981].

(v) If one expands the canonical quantities (I,3) in powers of $(1/L)$, then one gets

$$p^{\mu}_a = \pi^{\mu}_a + \mathrm{O}(1/L), \qquad H_c = H_{\mathrm{DW}} + \mathrm{O}(1/L), \qquad H_{\mathrm{DW}} = \pi^{\mu}_a v^a_{\mu} - L,$$

which shows that the standard canonical framework used in physics can be obtained from Carathéodory's one as the zero order approximation of a polynomial expansion in the variable $(1/L)$!

These examples show that Carathéodory's canonical theory for fields has very interesting elements concerning the qualitative dynamical and geometrical aspects of a given field theory.

In a series of important papers the Belgian mathematician Lepage [1936a, b, 1941, 1942a, b] showed that the theories of DeDonder–Weyl and Carathéodory are just special cases in a general framework of possible canonical theories for systems which have at least 2 independent and 2 dependent variables. The backbone of Lepage's analysis is E. Cartan's geometrical interpretation of partial differential equations and his use of differential forms in this context: According to Cartan the solutions $z^a = f^a(x)$ of partial differential equations define m-dimensional submanifolds in an $(m + n)$-dimensional space. The partial differential equations constitute conditions on the tangent spaces of these submanifolds, conditions which conveniently can be expressed in terms of those differential forms which vanish on the tangent spaces of the submanifolds.

By consequently exploiting the properties of differential forms Lepage was able to identify essential features of the Legendre transformation and to find, in a sense, the most general canonical framework (canonical momenta, Hamilton function, HJ equation etc.) for a given field theory defined by a Lagrangian. The canonical framework usually employed in physics is that of DeDonder–Weyl. This framework is, however, unsatisfactory as far as the transversality properties of extremals and wave fronts are concerned (see above), a disease which does not exist in the algebraically more complicated theory of Carathéodory.

As Lepage's very rich and interesting canonical framework – including HJ theories for fields – is not generally known, it is the main purpose of this article to draw attention to it and to indicate by physical examples how this more general framework may become useful for physics.

Chapter 1 collects those essential properties of differential forms which are being used later. The most important concepts here are the "rank" and "class" of a differential p-form, because these properties are crucial for the dimension of the integral submanifolds associated with a given p-form.

Chapter 2 recalls – in the language to be used later – a number of concepts from mechanics which are to be generalized to field theories. This chapter is purely pedagogical.

Chapter 3 introduces the main ideas of Lepage in the case of 2 independent and n dependent variables. 2 independent variables suffice in order to discuss the essential features of Lepage's general canonical framework: The generalized Legendre transformation, the replacement of the second order Euler–Lagrange equations by first order canonical equations, the classification of different canonical theories according to the rank of the basic canonical 2-form etc. Several of these concepts are illustrated by an application to 2-dimensional E-dynamics.

Chapter 4 discusses the concept of HJ theories for fields, their integrability problems, the questions associated with the transversality properties of extremals and wave fronts, some of the main features of the DWHJ theory: conserved currents associated with a parameter-dependent solution $S^\mu(x, z; a)$, the concept of a complete integral and how to construct solutions of the field equations from it. Furthermore, the problem how to find transversal wave fronts for a given extremal or for an n-parameter family of extremals is treated.

Chapter 5 contains a detailed discussion of Carathéodory's canonical theory of fields. Illustrating examples are E-dynamics in 2 dimensions, the relativistic string and scalar field theories.

Chapter 6 indicates how the results in chapters 3–5 for 2 independent variables are to be generalized to m independent ones.

Chapter 7 discusses canonical properties, including HJ equations, for systems which are invariant under reparametrization: strings, membranes etc.

Chapter 8 deals with situations where the transversality properties of extremals and wave fronts become singular: caustics appear. It turns out that the solutions of the equations of motion which have a vanishing Lagrangian (density), $L = 0$, have a special mathematical and physical significance. This can be seen from examples in mechanics as well as in field theories.

The final chapter 9 lists a number of open problems – of which there are many – and tries to speculate, what directions of future research in this field may be useful and of interest for physics.

The present paper extends, improves and, in a few instances, corrects previous short communications on the same subject [Kastrup, 1977–82].

1. Some elements of differential geometry

This introductory chapter is intended to familiarize the interested reader with some of the basic notions of modern differential geometry, because these concepts are *by far* the most appropriate ones in order to describe the canonical theories for fields discussed in the following chapters. The reason is the following: The dynamics of a physical system with m independent variables x^μ, $\mu = 1, \ldots, m$, and n

dependent variables $z^a(x)$, $a = 1, \ldots, n$, is generally determined by a system of differential equations for the functions $z^a(x)$. In E. Cartan's geometrical interpretation of differential equations [E. Cartan, 1945; Kähler, 1934; Kuranishi 1957, 1967; Sternberg, 1964; Hermann, 1965; Godbillon, 1969; Ślebodziński, 1970; Dieudonné IV, 1974; Choquet-Bruhat et al., 1977; Estabrook, 1980] the solutions $z^a(x)$ define m-dimensional submanifolds in an $(l = m + n)$-dimensional space and the differential equations constitute conditions on the tangent spaces of these submanifolds, conditions which most conveniently can be expressed in terms of differential forms.

This geometrical approach towards the problem of solving ordinary and partial differential equations turns out to be extremely fruitful for generalizing the canonical formulation of mechanics to field theories, with some surprises in store!

Those who are already familiar with basic concepts of modern differential geometry (manifolds, vector fields, differential forms, exterior differentiation, interior product of forms and vector fields, Lie derivatives, Poincaré's lemma, Stokes' theorem, differential systems and ideals, Frobenius' integrability theorem, integral submanifolds, rank and class of a differential form of degree p etc.) can skip this chapter.

This is *not* a systematic introduction to the concepts just mentioned. Systematic treatments, containing proofs etc., can be found in the literature discussed in the bibliographic notes at the end of this chapter and in the text mentioned above.

In order to keep the following account as simple as possible I shall mainly use "local" formulations of the notions to be discussed, i.e. I shall use the language of local coordinate frames.

1.1. Manifolds and their tangent vectors

Loosely speaking, an l-dimensional manifold M^l with points p is locally in one-to-one correspondence to a neighbourhood U of an Euclidean space R^l with coordinates $y^\lambda(p)$, $\lambda = 1, \ldots, l$. If we cover the manifold with an "atlas" of fixed (partially overlapping) coordinate systems, we can express most properties of M^l in terms of local coordinates. M^l is "differentiable", if on any overlap $\mathrm{U} \cap \hat{\mathrm{U}} \neq \emptyset$ the coordinate change $y(p) \to \hat{y}(p)$ is differentiable.

Let $\mathrm{C}(t)\colon \mathrm{I} \to \mathrm{M}^l$, $\mathrm{I} = [t_1, t_2]$, $t_2 > t_1$, be a smooth curve $\{p(t)\} \subset \mathrm{M}^l$, $y(t) = y(p(t))$ and $\{f(y)\}$ a set of smooth real-valued (test-) functions, then the map $\mathrm{C}'(t)\colon \{f\} \to \mathrm{R}$, defined by

$$[\mathrm{C}'(t)](f) := \frac{\mathrm{d}f(y(t))}{\mathrm{d}t} = \frac{\mathrm{d}y^\lambda}{\mathrm{d}t}\, \partial_\lambda f_{|y=y(t)}\,, \qquad \partial_\lambda := \partial/\partial y^\lambda\,, \tag{1,1}$$

is called a *tangent vector* to C at $y(t)$. If we take the special curves $\mathrm{C}_\lambda\colon t = y^\lambda$, then $\mathrm{C}'_\lambda(t) = \partial_\lambda$ and we see from eq. (1,1) that the differential operator $e_\lambda = \partial_\lambda$, $\lambda = 1, \ldots, l$, can be interpreted as a basis of an l-dimensional vector space $\mathrm{T}_{y(p)}(\mathrm{M}^l)$ of tangent vectors at $p \in \mathrm{M}^l$, or, locally, at $y(p)$. The union $\cup_{p\in\mathrm{M}^l}\mathrm{T}_{y(p)} = \mathrm{T}(\mathrm{M}^l)$ of all tangent spaces is called the tangent bundle of M^l. A *vector field* $Y = a^\lambda(y)\partial_\lambda$ is a mapping which assigns to each point $p \in \mathrm{M}^l$ a tangent vector $Y_p \in \mathrm{T}_{y(p)}$. The vector field Y is continuous, smooth etc. if the functions $a^\lambda(y)$ are continuous, smooth etc. An integral curve $\mathrm{C}_t(Y)$ of the vector field Y at $y(p_0) = y_0$ is a curve with coordinates $y(t)$ such that

$$\mathrm{d}y^\lambda/\mathrm{d}t = a^\lambda(y)\,, \qquad y^\lambda(0) = y_0^\lambda\,, \qquad \lambda = 1, \ldots, l\,. \tag{1,2}$$

Thus, each vector field on M^l defines a system of ordinary differential equations (1,2) of first order. It follows from the theory of differential equations that locally there are unique solutions $\varphi_t(y_0) = y(t)$, $\varphi_{t=0}(y_0) = y_0$ for any given y_0 and the solutions $\varphi_t(y_0)$ can be interpreted as a 1-parameter group of local transformations of M^l into itself:

$$y_0 \to y_t = \varphi_t(y_0)\,, \qquad \varphi_{t+s}(y_0) = \varphi_s[\varphi_t(y_0)]\,, \qquad \varphi_0(y_0) = y_0\,. \tag{1,3}$$

The vector field Y is said to define a "*flow*" $\varphi_t(y_0)$ on M^l.
Example:
For a mechanical system with phase space $\mathrm{M}^l = \mathrm{P}^{2n}$, coordinates $y = (q^1, \dots, q^n, p_1, \dots, p_n)$ and Hamilton function $H(q,p)$ the integral curves $\varphi_t(q_0, p^0)$ of the vector field

$$Y_H = \frac{\partial H}{\partial p_j}\frac{\partial}{\partial q^j} - \frac{\partial H}{\partial q^j}\frac{\partial}{\partial p_j}$$

define the Hamiltonian "flow" of the system through the (initial) point $(q_0, p^0) \in \mathrm{P}^{2n}$.

On the other hand: Each 1-parameter local transformation group φ_t: $y \to \varphi_t(y)$ "induces" a vector field Y on M^l by

$$Yf(y) = \lim_{t\to 0} \frac{1}{t}\,[f(\varphi_t(y)) - f(y)]\,. \tag{1,4}$$

Example:
The dilatations $y \to e^t y$ induce the vector field $Y = y^\lambda \partial_\lambda$.

If $Y_{(1)} = a^\lambda_{(1)}(y)\partial_\lambda$, $Y_{(2)} = a^\lambda_{(2)}(y)\partial_\lambda$, are two vector fields, then their commutator

$$[Y_{(1)}, Y_{(2)}] = (a^\lambda_{(1)}\partial_\lambda a^\kappa_{(2)} - a^\lambda_{(2)}\partial_\lambda a^\kappa_{(1)})\partial_\kappa \tag{1,5}$$

is again a vector field.

1.2. Cotangent vectors and differential forms

Let $\mathrm{T}^*_y(\mathrm{M}^l)$ be the vector space dual to the tangent space T_y, i.e. T^*_y consists of all linear mappings of T_y into the real numbers. The basis of T^*_y dual to ∂_κ is denoted by $\mathrm{d}y^\lambda$, i.e. $\mathrm{d}y^\lambda(\partial_\kappa) = \delta^\lambda_\kappa$ (=Kronecker's symbol) and the elements of T^*_y are called "*cotangent vectors*". The union $\mathrm{T}^*(\mathrm{M}^l)$ of all cotangent spaces T^*_y forms the cotangent bundle.

The objects "dual" to the vector fields are the "*1-forms*" or "*Pfaffian forms*" $\omega = b_\lambda(y)\,\mathrm{d}y^\lambda$ which assign to each y an element of T^*_y and which, applied to a vector field $Y = a^\lambda(y)\,\partial_\lambda$, give the real function $\omega(Y) = b_\lambda(y)\,a^\lambda(y)$.

A differential form of degree p or, briefly, a *p-form* $\omega^p \in \Lambda^p$ is a multilinear mapping of degree p which maps any p vector fields $Y_1, \dots, Y_p$ at y into the real numbers, $\omega^p(Y_1, \dots, Y_p) \in \mathrm{R}$ and which is completely antisymmetric in its arguments, e.g. $\omega^2(Y_1, Y_2) = -\omega^2(Y_2, Y_1)$. (The letter "$p$" here has, of course, nothing to do with the points p of the manifold M^l, for which we used it before and which we shall denote by y from now on.)

The (linear) space Λ^p of all p-forms ω^p at y can be interpreted as the pth exterior power of the cotangent space T^*_y, where exterior multiplication is defined as follows:

If $\omega_1 \in \Lambda^p$ and $\omega_2 \in \Lambda^q$, their *exterior* ("*wedge*") *product* $\omega_1 \wedge \omega_2 \in \Lambda^{p+q}$ is given by

$$\omega_1 \wedge \omega_2(Y_1, \ldots, Y_{p+q}) = \frac{1}{p!q!} \sum_{\sigma \in S_{p+q}} (\operatorname{sgn} \sigma)\, \omega_1(Y_{\sigma(1)}, \ldots, Y_{\sigma(p)})\, \omega_2(Y_{\sigma(p+1)}, \ldots, Y_{\sigma(p+q)}), \tag{1,6}$$

where S_{p+q} is the set of all permutations σ of $(1, \ldots, p+q)$. The wedge product (1,6) has the properties

$$\omega_1 \wedge \omega_2 = (-1)^{pq}\, \omega_2 \wedge \omega_1, \qquad (f\omega_1) \wedge \omega_2 = f(\omega_1 \wedge \omega_2),$$
$$(\omega_1 \wedge \omega_2) \wedge \omega_3 = \omega_1 \wedge (\omega_2 \wedge \omega_3). \tag{1,7}$$

One has $\Lambda^1_y = T^*_y$ and if, e.g., $\omega_j = b^{(j)}_\lambda \, dy^\lambda$, $j = 1, 2$, then

$$\omega_1 \wedge \omega_2 = b^{(1)}_\lambda b^{(2)}_\kappa \, dy^\lambda \wedge dy^\kappa = \sum_{\lambda<\kappa} (b^{(1)}_\lambda b^{(2)}_\kappa - b^{(2)}_\lambda b^{(1)}_\kappa)\, dy^\lambda \wedge dy^\kappa$$
$$= \tfrac{1}{2}(b^{(1)}_\lambda b^{(2)}_\kappa - b^{(2)}_\lambda b^{(1)}_\kappa)\, dy^\lambda \wedge dy^\kappa.$$

A basis of Λ^p is $\{dy^{\lambda_1} \wedge \cdots \wedge dy^{\lambda_p}, \lambda_1 < \cdots < \lambda_p\}$ and therefore the dimension of Λ^p is $\binom{l}{p}$. An element ω^p of Λ^p has the representation

$$\omega^p = \sum_{\lambda_1 < \cdots < \lambda_p} \omega_{\lambda_1 \ldots \lambda_p}(y)\, dy^{\lambda_1} \wedge \cdots \wedge dy^{\lambda_p},$$

or, if the functions $\omega_{\lambda_1 \ldots \lambda_p}(y)$ are completely antisymmetric in their indices $\lambda_1, \ldots, \lambda_p$:

$$\omega^p = \frac{1}{p!}\, \omega_{\lambda_1 \ldots \lambda_p}(y)\, dy^{\lambda_1} \wedge \cdots \wedge dy^{\lambda_p}.$$

1.3. Maps of manifolds and of their tangent and cotangent spaces

Let N^k be a k-dimensional manifold with local coordinates u^κ, $\kappa = 1, \ldots, k$. A *map* φ: $M^l \to N^k$, $u^\kappa = \varphi^\kappa(y)$ induces a map of tangent spaces, such that for a function $g(u)$ at $u = \varphi(y)$ and for $V_y \in T_y(M^l)$:

$$\varphi_*: T_y(M^l) \to T_{u=\varphi(y)}(N^k), \tag{1,8}$$

or, with $V_y = v^\lambda \partial_\lambda$,

$$[\varphi_*(V_y)_{u=\varphi(y)}]\, g(u) = V_y[g(\varphi(y))].$$

If $Y = a^\lambda(y)\partial_\lambda$ is a vector field, it follows that

$$[\varphi_*(Y)_{u=\varphi(y)}]\, g(u) = a^\lambda(y)\, (\partial\varphi^\kappa/\partial y^\lambda)\, \partial g(u)/\partial u^\kappa. \tag{1,9}$$

The r.h. side of eq. (1,9) defines a vector field on N^k iff the coordinates y can be expressed as functions of u, i.e. if the map $\varphi: M^l \to N^k$ is one-to-one.

Whereas the tangent map φ_*: $T_y(M^l) \to T_{u=\varphi(y)}(N^k)$ has the same "direction" as φ itself, the map φ induces a map φ^* of cotangent spaces $T^*_{u=\varphi(y)}(N^k) \to T^*_y(M^l)$ which "pulls" forms on N^k "back" to M^l:

Let ω be an element of $\Lambda^p_{u=\varphi(y)}(N^k)$, then $(\varphi^*\omega)_y \in \Lambda^p_y(M^l)$ is defined by

$$(\varphi^*\omega)_y(Y_1, \ldots, Y_p) = \omega_{\varphi(y)}(\varphi_*(Y_1), \ldots, \varphi_*(Y_p)) . \tag{1,10}$$

For instance, if $\omega = \omega_\kappa(u)\,du^\kappa$, then $\varphi^*\omega = \omega_\kappa(u = \varphi(y))\partial_\lambda\varphi^\kappa\, dy^\lambda$. In addition one has $\varphi^*(\omega_1 \wedge \omega_2) = \varphi^*(\omega_1) \wedge \varphi^*(\omega_2)$. Notice that the cotangent map φ^* does not require the map φ to be one-to-one. In this sense differential forms behave more "decent" under maps than vector fields!

1.4. Operations on differential forms

The *differential* df *of a function* $f(y)$ is a 1-form with the property $df(Y) = Yf(y)$, $Y = a^\lambda(y)\,\partial_\lambda$. We have $df(\partial_\lambda) = \partial_\lambda f$, $dy^\kappa(\partial_\lambda) = \partial_\lambda y^\kappa = \delta^\kappa_\lambda$ and therefore $df = \partial_\lambda f\, dy^\lambda$.

The following operations with differential forms are important:

(i) *Exterior differentiation*:

This is a mapping d: $\Lambda^p \to \Lambda^{p+1}$, with the properties:

(a) If $f(y) \in \Lambda^0$ is a function, then df is the differential of f.

(b) If ω_1 is a p-form, then

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^p \omega_1 \wedge d\omega_2 .$$

(c) $d(d\omega) = 0$ for all forms ω.

In local coordinates: If

$$\omega = \frac{1}{p!}\,\omega_{\lambda_1\ldots\lambda_p}\, dy^{\lambda_1} \wedge \cdots \wedge dy^{\lambda_p} ,$$

then

$$d\omega = \frac{1}{p!}(d\omega_{\lambda_1\ldots\lambda_p}(y)) \wedge dy^{\lambda_1} \wedge \cdots \wedge dy^{\lambda_p} . \tag{1,11}$$

If ω is a 1-form, then

$$d\omega(Y_1, Y_2) = Y_1\,\omega(Y_2) - Y_2\,\omega(Y_1) - \omega([Y_1, Y_2]) . \tag{1,12}$$

The representation (1,11) shows that $d\varphi^*(\omega) = \varphi^*(d\omega)$, i.e. exterior differentiation commutes with mappings! (For a 0-form (=function) $g(u)$ one defines $(\varphi^*g)(y) = g(\varphi(y))$.)

If a p-form ω has the representation $\omega = d\Theta$, where Θ is a $(p-1)$-form, then ω is called "*exact*". If, on the other hand, $d\omega = 0$, ω is called "closed". In general closed forms are not exact! The difference plays an important role in algebraic topology [see, e.g., Singer and Thorpe, 1976].

Locally, however, one has *Poincaré's* famous *lemma*: If $d\omega = 0$ in a region $\subset M^l$ which is contractible into each point of that region, then there is a $(p-1)$-form Θ, such that $\omega = d\Theta$. The best known application in physics is the following: Let

$$F = \tfrac{1}{2}F_{\mu\nu}\, dx^\mu \wedge dx^\nu\,, \qquad F_{\nu\mu} = -F_{\mu\nu}$$

be the 2-form defined by the electromagnetic fields $(F_{01}, F_{02}, F_{03}) = (E^1, E^2, E^3) = \boldsymbol{E}(x)$, $(F_{23}, F_{31}, F_{12}) = -(B^1, B^2, B^3) = -\boldsymbol{B}(x)$. If $dF = 0$ in a contractible region of Minkowski space, i.e. if the homogeneous Maxwell equations hold there, then, according to Poincaré's lemma, $F = dA$, $A = A_\mu(x)\, dx^\mu$. In the Aharonov–Bohm experiment, however, not all regions in question are contractible and the potential form A of F can be defined only locally [Aharonov and Bohm, 1959; Wu and Yang, 1975].

(ii) *Interior multiplication of a form by a vector field* Y:

This operation is a mapping $i(Y)$: $\Lambda^p \to \Lambda^{p-1}$, defined by

$$\begin{aligned} &i(Y) f = 0, \qquad i(Y)\, df = Yf, \\ &i(Y)(\omega_1 \wedge \omega_2) = (i(Y)\omega_1) \wedge \omega_2 + (-1)^p \omega_1 \wedge (i(Y)\omega_2)\,, \qquad \omega_1 \in \Lambda^p\,. \end{aligned}$$

If ω is a p-form, then $i(Y)\omega$ is a $(p-1)$-form

$$[i(Y)\,\omega](Y_1, \ldots, Y_{p-1}) = \omega(Y, Y_1, \ldots, Y_{p-1})\,. \tag{1,13}$$

Example:

$$\begin{aligned} i(\partial_\lambda)F &= \tfrac{1}{2}F_{\mu\nu}(dx^\mu(\partial_\lambda)\, dx^\nu - dx^\mu\, dx^\nu(\partial_\lambda)) \\ &= F_{\lambda\nu}\, dx^\nu\,. \end{aligned}$$

(iii) *Lie derivative of a form*:

Whereas exterior differentiation d maps Λ^p into Λ^{p+1} and interior multiplication $i(Y)$ maps Λ^p into Λ^{p-1}, the Lie derivative $L(Y)$ of forms with respect to a vector field Y,

$$L(Y) = i(Y)\, d + d\, i(Y) \tag{1,14}$$

maps Λ^p into Λ^p. $L(Y)$ has the properties

$$L(Y) f = Yf\,, \qquad L(Y)\, df = d(Yf)\,.$$

As Λ^p can be generated locally by functions f and their differentials df, the following properties of $L(Y)$ can be proven by applications to f and df

$$\begin{aligned} &L(Y)\, d = dL(Y)\,, \qquad L(Y_1 + Y_2) = L(Y_1) + L(Y_2)\,, \\ &L(fY) = f L(Y) + df \wedge i(Y)\,, \qquad [L(Y_1), L(Y_2)] = L([Y_1, Y_2]) \\ &[L(Y_1), i(Y_2)] = i([Y_1, Y_2])\,. \end{aligned} \tag{1,15}$$

If

$$\omega = \frac{1}{p!}\, \omega_{\lambda_1 \ldots \lambda_p}(y)\, dy^{\lambda_1} \wedge \cdots \wedge dy^{\lambda_p}$$

then

$$L(\partial_\mu)\,\omega = \frac{1}{p!}(\partial_\mu \omega_{\lambda_1\ldots\lambda_p}(y))\,\mathrm{d}y^{\lambda_1} \wedge \cdots \wedge \mathrm{d}y^{\lambda_p}\,.$$

The Lie derivative $L(Y)$ of a form ω with respect to a vector field Y can be defined in a different way: Let φ_t be the 1-parameter local transformation group generated by Y, then

$$L(Y)\,\omega = \lim_{t\to 0}\frac{1}{t}(\varphi_t^*\omega - \omega)\,. \tag{1,16}$$

In order to prove the equivalence of the definitions (1,14) and (1,16) it suffices – see the remark above – to apply both to functions f and their differentials $\mathrm{d}f$:

$$L(Y)f = Yf = \lim_{t\to 0}\frac{1}{t}(\varphi_t^* f - f) = \lim_{t\to 0}\frac{1}{t}[f(\varphi_t(y)) - f(y)]\,,$$

$$L(Y)\,\mathrm{d}f = \mathrm{d}(Yf) = \mathrm{d}\left[\lim_{t\to 0}\frac{1}{t}(\varphi_t^* f - f)\right] = \lim_{t\to 0}\frac{1}{t}[\varphi_t^*(\mathrm{d}f) - \mathrm{d}f]\,.$$

A p-form ω is said to be invariant with respect to the vector field Y if $L(Y)\omega = 0$. The meaning of this is obvious from the definition (1,16).

1.5. Stokes' theorem

We next turn to the very elegant formulation of Stokes' integral theorem in terms of differential forms: Let ω be a p-form with continuous coefficients in a region $\mathrm{G} \subset \mathrm{M}^l$ with boundary $\partial\mathrm{G}$, then the integral

$$\int_{\mathrm{G}} \omega \quad \text{is defined by} \sum_{\lambda_1 < \cdots < \lambda_p} \int_{\mathrm{G}} \omega_{\lambda_1\ldots\lambda_p}(y)\,\mathrm{d}y^{\lambda_1}\cdots\mathrm{d}y^{\lambda_p}\,.$$

If $\varphi: \mathrm{M}^l \to \mathrm{N}^k$ is a mapping which maps G on $\varphi(\mathrm{G})$, ω a p-form on $\varphi(\mathrm{G})$ and $\varphi^*\omega$ its "pull-back" on G, then

$$\int_{\varphi(\mathrm{G})} \omega = \int_{\mathrm{G}} \varphi^*\omega\,. \tag{1,17}$$

Stokes' theorem can now be stated as follows: If G is a p-dimensional region $\subset \mathrm{M}^l$ and ω a $(p-1)$-form, then

$$\int_{\mathrm{G}} \mathrm{d}\omega = \int_{\partial\mathrm{G}} \omega\,. \tag{1,18}$$

In order that eq. (1,18) makes sense, the region G, its boundary ∂G and the form ω have to fulfill certain conditions, a detailed discussion of which can be found in the literature mentioned at the end of this chapter.

An important application of eqs. (1,14), (1,16) and (1,18) in the following chapters will be the following: Let $\varphi_\tau(y)$, $\tau \in [0, 1]$, be a 1-parameter "variation" of M^l: $\varphi_\tau(y) \in \mathrm{M}^l$, $\varphi_{\tau=0}(y) = y$ and let V be the vector field on M^l which is induced by this variation, then according to eq. (1,17):

$$A_\tau := \int_{\varphi_\tau(\mathrm{G})} \omega = \int_{\mathrm{G}} \varphi_\tau^* \omega$$

and

$$\left.\frac{\mathrm{d}A_\tau}{\mathrm{d}\tau}\right|_{\tau=0} = \lim_{\tau\to 0} \frac{1}{\tau}(A_\tau - A_0) = \int_{\mathrm{G}} \lim_{\tau\to 0} \frac{1}{\tau}(\varphi_\tau^*\omega - \omega) = \int_{\mathrm{G}} L(V)\,\omega\,.$$

Combining this with eqs. (1,14) and (1,18) we get

$$\left.\frac{\mathrm{d}A_\tau}{\mathrm{d}\tau}\right|_{\tau=0} = \int_{\mathrm{G}} L(V)\,\omega = \int_{\mathrm{G}} i(V)\,\mathrm{d}\omega + \int_{\partial\mathrm{G}} i(V)\,\omega\,. \tag{1,19}$$

Eq. (1,19) is a basic formula in the calculus of variation, where, in conventional language, the notation $\delta y^\lambda = (\partial\varphi^\lambda(y)/\partial\tau)_{\tau=0}\tau + \cdots$ is used, instead of vector fields.

1.6. Vector fields, differential ideals, their integral submanifolds and Frobenius' integrability criterion

We have seen that a vector field $Y = a^\lambda(y)\,\partial_\lambda$ on a manifold M^l defines a system of ordinary differential eqs. (1,2), the solutions $y(t) = \varphi(t)$ of which are integral curves of Y in M^l, in other words: *One* vector field leads to 1-dimensional "integral" submanifolds $\mathrm{I}^1(Y) = \{y(t),\ y(0) = y_0\}$ through each point y_0 of M^l.

Correspondingly, several linearly independent vector fields $X_1, \ldots, X_m$ define a system of partial differential equations of first order, the solutions of which define m-dimensional integral submanifolds $\mathrm{I}^m(X_1, \ldots, X_m) = \{y = y(x),\ x = (x^1, \ldots, x^m)\}$, *if* certain *integrability conditions* are fulfilled. This last requirement is a new, but important one. It comes about as follows:

Locally one can choose a coordinate system $y = (x^1, \ldots, x^m, z^1, \ldots, z^n)$, $m + n = l$, such that

$$X_\mu = \partial_\mu + \varphi_\mu^a(x, z)\,\partial_a\,, \qquad \mu = 1, \ldots, m\,,$$

$$\partial_\mu := \partial/\partial x^\mu\,, \qquad \partial_a := \partial/\partial z^a\,. \tag{1,20}$$

In this coordinate system the integral manifolds $\mathrm{I}^m(X_1, \ldots, X_m)$ – if they exist – are given by the functions $z^a = f^a(x)$, $a = 1, \ldots, n$, which have to obey the partial differential equations

$$\partial_\mu f^a(x) = \varphi_\mu^a(x, z)\,. \tag{1,21}$$

This follows from eq. (1,1), if we put $t = x^\mu$, $\mu = 1, \ldots, m$. Because $\partial_\mu\partial_\nu f^a(x) = \partial_\nu\partial_\mu f^a(x)$ the functions

$\varphi^a_\mu(x, z)$ have to obey the integrability conditions

$$\frac{\mathrm{d}}{\mathrm{d}x^\nu}\varphi^a_\mu(x, f(x)) = \partial_\nu\varphi^a_\mu + \partial_b\varphi^a_\mu\varphi^b_\nu = \frac{\mathrm{d}}{\mathrm{d}x^\mu}\varphi^a_\nu(x, f(x)) = \partial_\mu\varphi^a_\nu + \partial_b\varphi^a_\nu\varphi^b_\mu . \tag{1,22}$$

In terms of the vector fields (1,20) the conditions (1,22) mean

$$[X_\mu, X_\nu] = 0 , \qquad \mu, \nu = 1, \ldots, m .$$

The last equations form a special case of *Frobenius' integrability criterion*, which asserts when a system $\mathscr{X}^m$ of m-dimensional tangent subspaces on M^l, the basis of which at each point y is given by m linearly independent smooth vector fields $Y_1, \ldots, Y_m$, form the tangent spaces of an m-dimensional integral submanifold $\mathrm{I}^m(Y_1, \ldots, Y_m) \subset \mathrm{M}^l$, associated with the "*differential system*" $\mathscr{X}^m$. The criterion says that the submanifolds I^m exist iff for $Y_\mu, Y_\nu \in \mathscr{X}^m$ the commutator $[Y_\mu, Y_\nu]$ belongs to $\mathscr{X}^m$, too. An application of this criterion is well-known: A subspace of a Lie-algebra generates a subgroup (submanifold of the whole group manifold) iff that subspace forms a Lie subalgebra!

According to E. Cartan it is more advantageous not to use the m vector fields (1,20), but those n Pfaffian forms ω^a, $a = 1, \ldots, n$, which "annihilate" the vector fields, i.e. for which $\omega^a(X_\mu) = 0$ for all a and μ. This is the case iff $\omega^a = \mathrm{d}z^a - \varphi^a_\mu\,\mathrm{d}x^\mu$.

The forms ω^a generate an ideal $\mathrm{I}[\omega^a]$ in the algebra $\Lambda = \Lambda^0 \oplus \Lambda^1 \oplus \cdots \oplus \Lambda^l$ of forms on M^l: If ω^a vanishes on $\mathscr{X}^m$, so does $\omega \wedge \omega^a$, where ω is an arbitrary element of Λ. The integrability criterion $[Y_\mu, Y_\nu] \in \mathscr{X}^m$, if $Y_\mu, Y_\nu \in \mathscr{X}^m$, is equivalent to the requirement that $\mathrm{d}\omega \in \mathrm{I}[\omega^a]$ if $\omega \in \mathrm{I}[\omega^a]$. This follows immediately from eq. (1,12). The ideal $\mathrm{I}[\omega^a]$ is called a *Pfaffian differential ideal.* More generally: If $\mathrm{I}[\omega]$ is an ideal of a set $\{\omega\}$ of differential forms – not only 1-forms –, then $\mathrm{I}[\omega]$ is called a *differential ideal* if $\mathrm{d}\omega \in \mathrm{I}[\omega]$ for all $\omega \in \mathrm{I}[\omega]$. For further details see, e.g., ch. IV.C of Choquet-Bruhat et al. [1977] or ch. XVIII of Dieudonné IV [1974]!

A consequence of Frobenius' integrability criterium is: If the ideal I is completely integrable then one can introduce local coordinates $\bar{y}^\lambda$, $\lambda = 1, \ldots, l$, such that the ideal $\mathrm{I}[\omega^a]$ is generated by the differentials $\mathrm{d}\bar{y}^{m+a}$, $a = 1, \ldots, n$ and the integral submanifolds I^m are characterized by the equations $\bar{y}^{m+a} = \text{const.}$, $a = 1, \ldots, n$.

1.7. Rank and class of a differential p-form

We now come to the most important part of this chapter: Let ω be a p-form on M^l. The minimal number r of linearly independent 1-forms $\theta^{(\rho)} = f^{(\rho)}_\lambda(y)\,\mathrm{d}y^\lambda$, $\rho = 1, \ldots, r$, by which ω can be expressed, is called the *rank* of ω. Obviously one has $r \geq p$; if $r = p$, then the form ω is called "simple" or "decomposable", in which case $\omega = \theta^{(1)} \wedge \cdots \wedge \theta^{(p)}$.

At each point y the forms $\theta^{(\rho)}$ generate an r-dimensional subspace of the cotangent space T^*_y. The union of these subspaces generated by the forms $\theta^{(\rho)}$ is called the "*associated system*" $\mathrm{A}^*(\omega)$ of ω.

The Pfaffian system $\mathrm{A}^*(\omega)$ determines an $(l - r)$-dimensional differential system $\mathrm{A}(\omega)$ of vector fields $Y = a^\lambda(y)\,\partial_\lambda$ which are solutions of the equations

$$\theta^{(\rho)}(Y) = f^{(\rho)}_\lambda a^\lambda = 0 , \qquad \rho = 1, \ldots, r . \tag{1,23}$$

The space $\mathrm{A}(\omega)$ is called the "*associated subspace*" ($\subset \mathrm{T}(\mathrm{M}^l)$) or "*associated differential system*" of the form ω. Its elements Y are the associated vector fields of ω. A vector field Y is associated with the form ω iff $i(Y)\,\omega = 0$. This follows from the fact that ω can be expressed by the 1-forms $\theta^{(\rho)}$ and that Y obeys the eqs. (1,23). A consequence is that the Pfaffian system $\mathrm{A}^*(\omega)$ is generated by the 1-forms

$$i(\partial_{\lambda_{p-1}})\cdots i(\partial_{\lambda_1})\omega\,,\qquad \lambda_1,\ldots,\lambda_{p-1}=1,\ldots,l\,.$$

The reason is that, because of $i(Y_1)\,i(Y_2) = -i(Y_2)\,i(Y_1)$,

$$\begin{aligned} i(Y)\,[i(\partial_{\lambda_{p-1}})\cdots i(\partial_{\lambda_1})\omega] &= (-1)^{p-1}\, i(\partial_{\lambda_{p-1}})\cdots i(\partial_{\lambda_1})\, i(Y)\omega \\ &= 0\,. \end{aligned}$$

Example:
For a 2-form $\omega = \frac{1}{2}\omega_{\kappa\lambda}\,\mathrm{d}y^\kappa \wedge \mathrm{d}y^\lambda$ we get $i(\partial_\kappa)\omega = \omega_{\kappa\lambda}\,\mathrm{d}y^\lambda$ which shows that the rank of ω is the same as that of the matrix $(\omega_{\kappa\lambda})$. As the rank of a skew symmetric matrix is always even, the rank r of any 2-form is even. Thus $r < l$ for uneven l. As an application consider the 2-form

$$F = \tfrac{1}{2}F_{\mu\nu}\,\mathrm{d}x^\mu \wedge \mathrm{d}x^\nu$$

of electromagnetic fields. In general F will have rank 4 and the associated system $\mathrm{A}^*(F)$ can be spanned by the four 1-forms $\mathrm{d}x^\mu$, $\mu = 0, 1, 2, 3$, and the only associated vector field X is the trivial one $X = 0$. The reason is that the equations

$$i(X)\,F = a^\lambda F_{\lambda\nu}\,\mathrm{d}x^\nu = 0\,,$$

or

$$a^\lambda F_{\lambda\nu} = 0\,,\qquad \nu = 0,\ldots,3\,,$$

have only the trivial solution $X = 0$ iff $\det(F_{\mu\nu}) = (\boldsymbol{E}\cdot\boldsymbol{B})^2 \neq 0$. If, however, $\det(F_{\mu\nu}) = 0$ and $(F_{\mu\nu}) \neq 0$, then F has rank 2 and its associated differential system $\mathrm{A}(F)$ can be spanned by the two vector fields [Choquet-Bruhat et al., 1977, p. 251]:

$$X_1 = \sum_{j=1}^{3} B^j\partial_j\,,\qquad X_2 = \boldsymbol{B}^2\partial_0 - \sum_{j=1}^{3}(\boldsymbol{E}\times\boldsymbol{B})^j\partial_j\,. \tag{1,24}$$

In general the differential system $\mathrm{A}(\omega)$ (or the Pfaffian system $\mathrm{A}^*(\omega)$) is not completely integrable, i.e. in general there is no $(l-r)$-dimensional integral submanifold $\mathrm{I}^{(l-r)}[\mathrm{A}(\omega)] = \mathrm{I}^{(l-r)}(\omega)$ the tangent spaces $\mathrm{T}_y(\mathrm{I}^{(l-r)})$ of which at each point y are spanned by $(l-r)$ linearly independent vector fields Y_σ, $\sigma = 1,\ldots,l-r$, which belong to $\mathrm{A}(\omega)$.

A completely integrable differential system $\mathrm{C}(\omega)$ associated with a form ω is obtained as follows: Let $\mathrm{C}(\omega)$ be the intersection of the differential systems $\mathrm{A}(\omega)$ and $\mathrm{A}(\mathrm{d}\omega)$. The subspace $\mathrm{C}(\omega)$ is called the *characteristic subspace* or the *characteristic differential system* of ω. It can be shown that $\mathrm{C}(\omega)$ is the largest completely integrable subspace of $\mathrm{A}(\omega)$. It follows immediately that the space $\mathrm{C}^*(\omega)$ of Pfaffian forms which annihilate $\mathrm{C}(\omega)$ is the union of the Pfaffian systems $\mathrm{A}^*(\omega)$ and $\mathrm{A}^*(\mathrm{d}\omega)$. $\mathrm{C}^*(\omega)$ is called the *characteristic Pfaffian system* of ω. The dimension c of the characteristic system $C^*(\omega)$ is called the "*class*" of ω. One sees that class $c = \text{rank } r$ if $d\omega = 0$ and $c \geq r$ in general. A vector field Y is an element of $C(\omega)$ iff

$$i(Y)\,\omega = 0\,, \qquad i(Y)\,d\omega = 0\,. \tag{1,25}$$

Because of eq. (1,14) the eqs. (1,25) are equivalent to

$$i(Y)\,\omega = 0\,, \qquad L(Y)\,\omega = 0\,. \tag{1,26}$$

It then follows from the last two of eqs. (1,15) that the commutator $[Y_1, Y_2]$ of two characteristic vector fields is again a characteristic vector field:

$$i([Y_1, Y_2])\omega = L(Y_1)\,i(Y_2)\,\omega - i(Y_2)\,L(Y_1)\,\omega = 0\,,$$

$$L([Y_1, Y_2])\,\omega = L(Y_1)\,L(Y_2)\,\omega - L(Y_2)\,L(Y_1)\,\omega = 0\,.$$

This shows that the characteristic subspace $C(\omega)$ is completely integrable. The corresponding characteristic integral submanifolds are $(l-c)$-dimensional. Consider now the largest subspace of $A(\omega)$ which is closed under commutator building, then it follows from eq. (1,12) that this subspace will be annihilated by $A^*(\omega)$ and $A^*(d\omega)$. Therefore it must be contained in $C(\omega)$ and because of the properties of $C(\omega)$ the two are the same.

It follows from Frobenius' integrability criterion that for a given p-form there exist c functions $f^{(1)}(y), \ldots, f^{(c)}(y)$ such that the characteristic Pfaffian system $C^*(\omega)$ is generated by the differentials $df^{(1)}, \ldots, df^{(c)}$ and the $(l-c)$-dimensional characteristic integral submanifolds $I^{(l-c)}[C(\omega)]$ are given by $f^{(1)}(y) = \text{const.}, \ldots, f^{(c)}(y) = \text{const.}$!

Examples:

(i) Suppose the electromagnetic field form F has rank 2. Because it is a closed form, $dF = 0$, its class c is 2, too. That means that the differential system (1,24) is completely integrable and defines 2-dimensional submanifolds $S^j(x) = \text{const.}$, $j = 1, 2$, of the Minkowski space M^4. According to our discussion above, the functions $S^j(x)$ obey the relation

$$F = dS^1 \wedge dS^2\,, \qquad F_{\mu\nu} = \partial_\mu S^1\, \partial_\nu S^2 - \partial_\mu S^2\, \partial_\nu S^1\,. \tag{1,27}$$

As $dS^1 \wedge dS^2 = d(S^1\, dS^2) = -d(S^2\, dS^1)$ we can interpret

$$S^1\, dS^2 = S^1\, \partial_\mu S^2\, dx^\mu\,, \qquad \text{or} \qquad -S^2\, dS^1 \tag{1,28}$$

as the "potential" 1-form of an electromagnetic field of rank 2. A gauge transformation is performed by adding the differential dg of a function $g(x)$ to the potential forms (1,28). (The choice $g = -S^1 S^2$ transforms the first one into the second one!).

The submanifolds $S^j(x) = \text{const.}$, $j = 1, 2$, can be interpreted as the 2-dimensional Hamilton–Jacobi "wave fronts" associated with the motion of a relativistic string [Rinke, 1980; Kastrup and Rinke, 1981].

(ii) Consider on $G^{n+1} = \{y = (t, q^1, \dots, q^n)\} \subset R^{n+1}$ the 1-form

$$\omega = -H\,dt + \psi_j\,dq^j\,,$$
$$\psi_j = \psi_j(t,q)\,,\qquad H = (t, q, \psi(t,q))\,. \tag{1,29}$$

The elements $w = w_t\partial_t + w^j\partial_j$ of the associated subspace $A(\omega)$ have to obey the equation

$$\omega(w) = -Hw_t + \psi_j w^j = 0\,.$$

If $H \neq 0$ this equation has the n independent solutions

$$w_{(j)} = \psi_j\partial_t + H\partial_j\,,\qquad j = 1, \dots, n\,. \tag{1,30}$$

In the following we shall use the notations

$$\partial_k H := \left.\frac{\partial H(t,q,\psi)}{\partial q^k}\right|_{\psi\ \text{fixed}}\,,\qquad D_k H := \partial_k H + \frac{\partial H}{\partial \psi_j}\partial_k\psi_j\,.$$

In general the vector fields (1,30) will not constitute a completely integrable system. For the commutator of any two of them we obtain

$$[w_{(j)}, w_{(k)}] = (HD_jH + \psi_j\partial_t H)w_{(k)} - (HD_kH + \psi_k\partial_t H)w_{(j)}$$
$$+ [H(\partial_j\psi_k - \partial_k\psi_j) + \psi_j(\partial_t\psi_k + D_kH) - \psi_k(\partial_t\psi_j + D_jH)]\partial_t\,. \tag{1,31}$$

The r.h. side of these equations is a linear combination of the vector fields $w_{(j)}(t,q)$ – i.e. $A(\omega)$ is completely integrable – if

$$\partial_t\psi_j + D_jH = 0\,,\qquad \partial_j\psi_k - \partial_k\psi_j = 0\,,\qquad j,k = 1,\dots,n\,. \tag{1,32}$$

If these relations are satisfied, then the form (1,29) is closed: $d\omega = 0$. Poincaré's lemma asserts that – at least locally – there exists a function $S(t,q)$ such that $dS(t,q) = \omega$, which yields a (Hamilton–Jacobi) equation for $S(t,q)$:

$$\partial_t S(t,q) = -H(t, q, \psi(t,q))\,,\qquad \partial_j S(t,q) = \psi_j(t,q)\,. \tag{1,33}$$

(iii) Consider the exterior derivative

$$d\theta = -dH(t,q,p) \wedge dt + dp_j \wedge dq^j$$
$$= -\left(\partial_j H\,dq^j + \frac{\partial H}{\partial p_j}dp_j + \partial_t H\,dt\right) \wedge dt + dp_j \wedge dq^j \tag{1,34}$$
$$= (dp_j + \partial_j H\,dt) \wedge \left(dq^j - \frac{\partial H}{\partial p_j}dt\right)$$

of the form $\theta = -H\,dt + p_j\,dq^j$. $d\theta$ is a form on R^{2n+1} with coordinates $y = (t, q^1, \ldots, q^n, p_1, \ldots, p_n)$. Its associated system – which coincides with its characteristic system, because $d\theta$ is exact – is generated by the $2n$ Pfaffian forms

$$i(\partial_j)\,d\theta = -(dp_j + \partial_j H\,dt) =: -\theta_j\,,$$

$$i(\partial/\partial p_j)\,d\theta = dq^j - \frac{\partial H}{\partial p_j}\,dt =: \omega^j\,. \tag{1,35}$$

If there are no additional constraints, rank and class of $d\theta$ are $2n$. The space $A(d\theta) = C(d\theta)$ of (characteristic) vector fields associated with $d\theta$ is 1-dimensional and can be generated by

$$X_H = \partial_t + \frac{\partial H}{\partial p_j}\partial_j - \frac{\partial H}{\partial q^j}\frac{\partial}{\partial p_j}\,,$$

because $\theta_j(X_H) = 0$, $\omega^j(X_H) = 0$. Thus, the associated integral manifolds of $d\theta$ are the 1-dimensional solutions $q^j(t)$, $p_j(t)$ of the differential equations

$$\dot q^j = \partial H/\partial p^j\,, \qquad \dot p_j = -\partial H/\partial q^j\,.$$

1.8. Some bibliographical notes (highly personal and selective)

The vivid freshness and intuitive directness of E. Cartan's monograph from 1945 is unsurpassed! An introduction to the concepts surveyed above which is very useful and appealing to physicists is chapter IV of Choquet-Bruhat et al. [1977]. Well written and equally helpful I found Godbillon [1969]. Much of the material necessary can also be found in Dieudonné's textbooks, vol. III [1972] and especially chapter XVIII of vol. IV [1974]. In addition I found the following more or less elementary introductions to modern differential geometry useful (they do not contain, however, the notions "rank" and "class" of a differential p-form): Pham Mau Quan [1969], Warner [1971], Matsushima [1972].

2. Reminiscences from mechanics

In order to prepare the ground for the application of Lepage's ideas to field theories and to illustrate some of the ideas involved in cases of familiar systems we first discuss some elements concerning the "canonical" framework of mechanical systems. We shall later see, however, that the full power of Lepage's ideas shows itself only, if we have systems with at least two independent and two dependent variables! Most of the literature concerned will be mentioned in the bibliographic notes at the end of this chapter.

2.1. The Legendre transformation $v^j \to p_j$, $L \to H$

Let $L = L(t, q, \dot q)$ be a (Lagrangian) function of the $2n+1$ variables $t \in [t_1, t_2]$, $q = (q^1, \ldots, q^n) \in G^n \subset R^n$ and $\dot q = (\dot q^1, \ldots, \dot q^n) \in R^n$. The time evolution of a mechanical system corresponds to a curve $C_0(t) = \{(t, q(t), \dot q(t) = dq(t)/dt)\}$ in $[t_1, t_2] \times G^n = G^{n+1}$. For a given system, corresponding to an appropriate choice of the function L, the "dynamical" or "real" curves $C_0(t)$ are characterized by the property that they make the action integral

$$A[C] = \int_{t_1}^{t_2} L(t, q, \dot{q})\, dt$$

stationary, if one compares the value of $A[C_0]$ with the values $A[C]$ of "neighboring" curves $C(t)$ – we shall make this notion more precise in a minute –. A familiar necessary condition is that the "extremals" $C_0(t)$ have to obey the Euler–Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}^j} - \partial_j L = 0\,, \qquad j = 1, \ldots, n\,. \tag{2,1}$$

The problem can be rephrased in the following way [Dixon, 1909; Kneser, 1921]: We consider $L = L(t, q, v)$ as a function of $2n$ equivalent dependent variables q and $v = (v^1, \ldots, v^n)$ with the condition (constraint) that $v^j = \dot{q}^j$ on the extremals. Such a problem can be treated with the help of Lagrangian multipliers λ_j: We consider the Lagrangian

$$\hat{L} = L(t, q, v) - \lambda_j(v^j - \dot{q}^j)\,.$$

The Euler–Lagrange equations for the variables v^j are

$$\frac{\partial \hat{L}}{\partial v^j} = \frac{\partial L}{\partial v^j} - \lambda_j = 0$$

because

$$\partial \hat{L}/\partial \dot{v}^j = 0$$

and we obtain

$$\hat{L}(t, q, \dot{q}, v) = L(t, q, v) - v^j\, \partial L/\partial v^j + \dot{q}^j\, \partial L/\partial v^j\,. \tag{2,2}$$

If we introduce $p_j = \partial L(t, q, v)/\partial v^j$ as new variables and assume the matrix

$$\left(\frac{\partial p_j}{\partial v^k}\right) = \left(\frac{\partial^2 L}{\partial v^k\, \partial v^j}\right) \tag{2,3}$$

to be regular, we can solve the equations $p_j = \partial L/\partial v^j$ for $v^j = \hat{\varphi}^j(t, q, p)$ and define

$$H(t, q, p) = \hat{\varphi}^j(t, q, p)\, p_j - L[t, q, v = \hat{\varphi}(t, q, p)]\,.$$

With these definitions we obtain for the Lagrangian (2,2):

$$\hat{L}(t, q, \dot{q}, p) = -H(t, q, p) + \dot{q}^j p_j\,. \tag{2,4}$$

This Lagrangian contains $2n$ dependent variables q and p, but it depends only on the derivatives of q, not on those of p. We therefore get as the $2n$ Euler–Lagrange equations for the $2n$ variables q and p:

$$-\frac{\partial \hat{L}}{\partial p_j} = \frac{\partial H}{\partial p_j} - \dot{q}^j = 0\,, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial \hat{L}}{\partial \dot{q}^j} - \frac{\partial \hat{L}}{\partial q^j} = \dot{p}_j + \partial_j H = 0\,. \tag{2,5}$$

The basic idea of Lepage is similar in spirit to the above introduction of Hamilton's function $H(t, q, p)$ and the derivation of the canonical equations of motion (2,5), but, due to the use of differential forms, it is much more efficient:

Let $\omega = L(t, q, \dot{q})\,\mathrm{d}t$ be the initial Lagrangian 1-form. The above procedure of introducing new variables v^j which are equal to $\dot{q}^j$ on the extremals is equivalent to introducing 1-forms $\omega^j = \mathrm{d}q^j - v^j\,\mathrm{d}t$ which vanish on the extremals $q(t)$, where $\mathrm{d}q^j(t) = \dot{q}^j\,\mathrm{d}t$, or, where the tangent vectors $e_t = \partial_t + \dot{q}^j\partial_j$ are annihilated by each ω^j:

$$\omega^j(e_t) = \dot{q}^j - v^j = 0\,, \qquad j = 1, \ldots, n\,.$$

Thus, as far as the extremals are concerned, the form $\omega = L\,\mathrm{d}t$ is only one representative in an equivalence class of 1-forms, the most general element of which is

$$\Omega = L(t, q, v)\,\mathrm{d}t + h_j\omega^j\,. \tag{2,6}$$

As discussed in chapter 1, the n Pfaffian forms ω^j generate an ideal $\mathrm{I}[\omega^j]$ which vanishes on the extremals.

The coefficients h_j – which correspond to the Lagrange parameters λ_j above – in general can be arbitrary functions of t, q and v. According to Lepage they can be determined by the following condition which has powerful generalizations in field theory:

For the exterior derivative of Ω we obtain

$$\begin{aligned}\mathrm{d}\Omega &= \left(\frac{\partial L}{\partial v^j} - h_j\right)\mathrm{d}v^j \wedge \mathrm{d}t + (-\partial_j L\,\mathrm{d}t + \mathrm{d}h_j) \wedge \omega^j \\ &= \left(\frac{\partial L}{\partial v^j} - h_j\right)\mathrm{d}v^j \wedge \mathrm{d}t + 0(\mathrm{mod}\ \mathrm{I}[\omega^j])\,.\end{aligned} \tag{2,7}$$

We therefore have $\mathrm{d}\Omega = 0$ on the extremals, where $\omega^j = 0$, iff $h_j = \partial L/\partial v^j =: p_j$. Inserting this value for h_j into Ω, we get

$$\begin{aligned}\Omega &= L\,\mathrm{d}t + p_j\omega^j = L\,\mathrm{d}t + p_j(\mathrm{d}q^j - v^j\,\mathrm{d}t) \\ &= -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j =: \theta\,.\end{aligned} \tag{2,8}$$

The eqs. (2,8) provide the following interpretation of the Legendre transformation

$$v^j \to p^j = \frac{\partial L}{\partial v^j}(t, q, v)\,, \qquad L \to H = v^j p_j - L:$$

According to our discussion above this transformation can be implemented by a change of basis

$$\mathrm{d}t \to \mathrm{d}t\,, \qquad \omega^j \to \mathrm{d}q^j \tag{2,9}$$

in the cotangent spaces $T^*_{(t,q)}(G^{n+1})$ of the region G^{n+1} on which the canonical form Ω is defined. Eqs. (2,8) show that the Hamilton function H is the resulting coefficient of $\mathrm{d}t$ and the canonical momenta p_j the resulting coefficients of $\mathrm{d}q^j$, after the change of basis (2,9) has been performed. This interpretation of the Legendre transformation turns out to be very powerful in field theories! In the following we assume the Legendre transformation to be regular.

The property: $\mathrm{d}\Omega = 0$ on the extremals is characteristic for the Hamilton–Jacobi theory of a mechanical system and the realizability of such a theory is one of Lepage's main motives.

2.2. Families of extremals and HJ theory

In a HJ theory one does not consider single extremals, but a family (a so-called "field") of them, such that through each point of a region $G^{n+1} = \{(t, q^1, , \ldots, q^n)\}$ passes just one extremal. Suppose the different extremals $q(t)$ are parametrized by constants (of integration) $u^1, \ldots, u^n, \ldots$ – there may be more than n parameters, but there should be at least n of them –, i.e. on G^{n+1} we have

$$q^j(t) = f^j(t; u)\,, \qquad p_j = g_j(t; u)\,, \qquad u = (u^1, \ldots, u^n, \ldots)\,.$$

The functions f^j and g_j obey the equations

$$\begin{aligned} \partial_t f^j(t; u) &= \frac{\partial H}{\partial p_j}[t, f(t; u), g(t; u)]\,, \\ \partial_t g_j(t; u) &= -\frac{\partial H}{\partial q^j}[t, f(t; u), g(t; u)]\,. \end{aligned} \tag{2,10}$$

If we define

$$\Lambda(t; u) := -H(t, f, g) + g_j \partial_t f^j \tag{2,11}$$

the function

$$\sigma(t, t_0; u) = \int_{t_0}^{t} \Lambda(\bar{t}; u)\,\mathrm{d}\bar{t} \tag{2,12}$$

has the properties:

$$\mathrm{d}\sigma = \Lambda(t; u)\,\mathrm{d}t - \Lambda(t_0; u)\,\mathrm{d}t_0 + \frac{\partial \sigma}{\partial u^k}\mathrm{d}u^k\,,$$

furthermore, because of

$$\frac{\partial \Lambda}{\partial u^k} = -\frac{\partial H}{\partial q^j}\frac{\partial f^j}{\partial u^k} - \frac{\partial H}{\partial p_j}\frac{\partial g_j}{\partial u^k} + \frac{\partial g_j}{\partial u^k}\partial_t f^j + g_j \partial_t \frac{\partial f^j}{\partial u^k} = \partial_t\left(g_j \frac{\partial f^j}{\partial u^k}\right),$$

where the eqs. (2,10) have been used, we obtain

$$\frac{\partial\sigma}{\partial u^k} = \int_{t_0}^{t} \mathrm{d}\bar{t}\, \partial\Lambda(\bar{t}, u)/\partial u^k = \left[g_j \frac{\partial f^j}{\partial u^k}\right]_{t_0}^{t} . \tag{2,13}$$

With

$$g_j\left(\partial_t f^j\, \mathrm{d}t + \frac{\partial f^j}{\partial u^k}\,\mathrm{d}u^k\right) = g_j\, \mathrm{d}f^j = p_j\, \mathrm{d}q^j$$

we get for the differential $\mathrm{d}\sigma(t, t_0; u)$:

$$\begin{aligned} &\mathrm{d}\sigma = -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j - (-H_0\,\mathrm{d}t_0 + p_j^{(0)}\,\mathrm{d}q^j_{(0)}) ,\\ &q^j_{(0)} = f^j(t_0; u) , \qquad p_j^{(0)} = g_j(t_0; u) , \qquad H_0 = H(t_0, q_{(0)}, p^{(0)}) \end{aligned} \tag{2,14}$$

or

$$\begin{aligned} -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j &= \mathrm{d}\sigma(t, t_0, u) - H_0\,\mathrm{d}t_0 + p_j^{(0)}\,\mathrm{d}q^j_{(0)} \\ &= \partial_t\sigma\,\mathrm{d}t + \left(\frac{\partial\sigma}{\partial u^k} + g_j(t_0; u)\frac{\partial f^j}{\partial u^k}(t_0, u)\right)\mathrm{d}u^k . \end{aligned} \tag{2,15}$$

The r.h. side of this equation is a differential of a function, iff

$$\frac{\partial}{\partial u^k}\,\partial_t\sigma = \partial_t\left[\frac{\partial\sigma}{\partial u^k} + g_j(t_0; u)\frac{\partial}{\partial u^k} f^j(t_0; u)\right] \tag{2,16a}$$

and

$$\frac{\partial}{\partial u^l}\left[\frac{\partial\sigma}{\partial u^k} + g_j(t_0; u)\frac{\partial}{\partial u^k} f^j(t_0; u)\right] = \frac{\partial}{\partial u^k}\left[\frac{\partial\sigma}{\partial u^l} + g_j(t_0; u)\frac{\partial}{\partial u^l} f^j(t_0; u)\right] . \tag{2,16b}$$

Since $g_j(t_0; u)(\partial f^j(t_0; u)/\partial u^k)$ is independent of t, the eqs. (2,16a) are automatically fulfilled. The eqs. (2,16b) are equivalent to

$$[u^k, u^l]_{t_0} := \left(\frac{\partial f^j}{\partial u^k}\frac{\partial g_j}{\partial u^l} - \frac{\partial f^j}{\partial u^l}\frac{\partial g_j}{\partial u^k}\right)_{t_0} = 0 . \tag{2,17}$$

The quantities $[u^k, u^l]$ are the so-called "Lagrange brackets". It follows from

$$\partial_t\left(\frac{\partial f^j}{\partial u^k}\frac{\partial g_j}{\partial u^l}\right) = \frac{\partial^2 H}{\partial p_i\,\partial p_j}\frac{\partial g_i}{\partial u^k}\frac{\partial g_j}{\partial u^l} - \frac{\partial^2 H}{\partial q^i \partial q^j}\frac{\partial f^i}{\partial u^k}\frac{\partial f^j}{\partial u^l}$$

that

$$\partial_t[u^k, u^l] = 0 \tag{2,18}$$

"along" a curve $q^j(t) = f^j(t; u)$, $p_j(t) = g_j(t; u)$, i.e. $[u^k, u^l]$ is constant along an extremal: If $[u^k, u^l] = 0$ at $t = t_0$, then it remains so for all t!

We therefore have the important result that

$$d\theta = d(-H\,dt + p_j\,dq^j) = 0 \qquad \text{iff } [u^k, u^l]_{t_0} = 0\,. \tag{2,19}$$

By an appropriate choice of the initial data at $t = t_0$: $f^j(t_0; u)$ and $g_j(t_0; u)$, we can – in general – always fulfill the condition $[u^k, u^l](t_0) = 0$, for instance, if $q(t_0; u) = 0$ for all u^j.

According to Poincaré's lemma – see section 1.4 – the property $d\theta = 0$ implies the existence of a function $S(t, q)$ such that the relation

$$dS(t, q) = -H\,dt + p_j\,dq^j \tag{2,20a}$$

holds, which is equivalent to the HJ equation

$$\partial_t S(t, q) + H(t, q, p) = 0\,, \qquad p_j = \partial_j S(t, q)\,. \tag{2,20b}$$

Suppose, there are n parameters u^k such that

$$\Delta(t; u) = \left| \left(\frac{\partial f^j}{\partial u^k}(t; u) \right) \right| \neq 0 \qquad \text{in } G^{n+1}\,, \tag{2,21}$$

then we can solve the n equations $q^j = f^j(t; u)$ for the parameters u^k:

$$u^j = \chi^j(t, q)\,, \qquad j = 1, \dots, n\,. \tag{2,22}$$

Inserting these functions into $f^j(t; u)$ and $g_j(t; u)$ we obtain

$$\dot{q}^j = \partial_t f^j(t; \chi(t, q)) =: \varphi^j(t, q)\,, \tag{2,23a}$$

$$p_j = g_j(t; \chi(t, q)) =: \psi_j(t, q)\,. \tag{2,23b}$$

The eqs. (2,23a) constitute a system of first-order differential equations for the extremals $q(t)$. They give the velocity of each "degree of freedom" at each point $(t, q) \in G^{n+1}$. For this reason the functions $\varphi^j(t, q)$ are said to define a "slope field".

The functions φ^j and ψ_j are related by a Legendre transformation: If the φ^j are given, we have

$$\psi_j(t, q) = \frac{\partial L}{\partial v^j}(t, q, v = \varphi(t, q))\,,$$

or, if we know the functions ψ_j, then

$$\varphi^j(t, q) = \frac{\partial H}{\partial p_j}(t, q, p = \psi(t, q))$$

provided the matrix (2,3) is regular.

Suppose now we are given n functions $\psi_j(t, q)$ and that we know a set of solutions $q^j(t) = f^j(t; u)$ of the differential equations

$$\dot{q}^j(t) = \frac{\partial H}{\partial p_j}[t, q, p = \psi(t, q)], \tag{2,24}$$

such that the inequality (2,21) holds. Since

$$\frac{\partial g_j}{\partial u^k} = \frac{\partial \psi_j}{\partial q^l}\frac{\partial f^l}{\partial u^k},$$

we have

$$[u^k, u^l] = \left(\frac{\partial \psi_j}{\partial q^i} - \frac{\partial \psi_i}{\partial q^j}\right)\frac{\partial f^j}{\partial u^k}\frac{\partial f^i}{\partial u^l}. \tag{2,25}$$

It follows that $[u^k, u^l] = 0$, if $\partial_i\psi_j = \partial_j\psi_i$. On the other hand, if $[u^k, u^l] = 0$, then $\partial_i\psi_j = \partial_j\psi_i$ because

$$\det\left(\frac{\partial f^j}{\partial u^k}\frac{\partial f^i}{\partial u^l}\right) = \Delta^{2n} \neq 0.$$

We here make direct contact with the problems discussed in chapter 1 in connection with the form (1,29): The integrability conditions (1,32) (second half) for the form

$$-H(t, q, \psi(t, q))\,\mathrm{d}t + \psi_j(t, q)\,\mathrm{d}q^j$$

are the same as above: $\partial_i\psi_j = \partial_j\psi_i$. The first half,

$$\partial_t\psi_j + \mathrm{D}_j H = 0 \tag{2,26}$$

of the conditions (1,32) in our context has the following interpretation: Suppose $q(t)$ is a curve in a region, where the functions $\psi_j(t, q)$ are defined. Then

$$\partial_t\psi_j = \frac{\mathrm{d}}{\mathrm{d}t}p_j - \partial_k\psi_j\dot{q}^k, \qquad p_j(t) = \psi_j(t, q(t)).$$

Since

$$\mathrm{D}_j H = \partial_j H + \frac{\partial H}{\partial \psi_k}\partial_j\psi_k,$$

the eqs. (2,26) give

$$\dot{p}_j + \partial_j H + \partial_k\psi_j(\partial H/\partial\psi_k - \dot{q}^k) = 0. \tag{2,27}$$

Thus, the integrability conditions (2,26) imply that the curves $q(t)$ are solutions of the equations

$\dot{p}_j + \partial_j H = 0$, if they are solutions of the eqs. (2,24). Notice that the integrability conditions are fulfilled automatically if $\psi_j = \partial_j S(t, q)$.
Example:
For $n = 1$ there are no nonvanishing Lagrange brackets. For $n = 2$ a simple illustration of the above discussion is provided by the harmonic oscillator in the plane with $L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2 - \omega^2(x^2 + y^2))$ and the extremals

$$x(t) = A_1 \sin \omega t + B_1 \cos \omega t\,, \qquad y(t) = A_2 \sin \omega t + B_2 \cos \omega t\,.$$

For any choice $(u^1, u^2) = (A_1, A_2)$, (A_1, B_2), (B_1, A_2) and (B_1, B_2) the Lagrange brackets $[u^1, u^2]$ vanish. Let us take $u^1 = B_1$, $u^2 = A_2$, $A_1 = 0$, $B_2 = 0$. We then have $\Delta(t{:}\, B_1, A_2) = \cos \omega t \sin \omega t$, which is $\neq 0$, provided $\omega t \neq n\pi/2$, $n = 0, \pm 1, \pm 2, \ldots$ (the points where $\Delta(t; u) = 0$ are called "focal points". They will be discussed in chapter 8). For the functions (2,22) we get here

$$B_1 = \chi^1(t, x, y) = x/\cos \omega t\,, \qquad A_2 = \chi^2(t, x, y) = y/\sin \omega t\,,$$

and the differential eqs. (2,23a) take the form

$$\begin{aligned} \dot{x} &= -B_1\omega \sin \omega t = -\omega x \,\mathrm{tg}\, \omega t = \varphi^1(t, x, y)\,, \\ \dot{y} &= A_2 \omega \cos \omega t = \omega y \,\mathrm{ctg}\, \omega t = \varphi^2(t, x, y)\,, \\ &\omega t \neq n\pi/2\,, \qquad n = 0, \pm 1, \pm \cdots . \end{aligned}$$

The functions $u^j = \chi^j(t, q)$ can be used for constructing a solution of the HJ eq. (2,20b) ("method of characteristics"): Let us assume the initial data at t_0 for the functions $q^j = f^j(t; u)$, $p_j = g_j(t; u)$ are such that $(g_j \partial f^j/\partial u^k)_{t_0} = 0$. Inserting $u^j = \chi^j(t, q)$ into the function $\sigma(t; u)$, eq. (2,12), then the resulting function $S(t, q) = \sigma(t; \chi(t, q))$ is a solution of the HJ equation.
Proof: With the help of the relation (2,13) we obtain

$$\begin{aligned} \partial_t S(t, q) &= \partial_t \sigma + \frac{\partial \sigma}{\partial u^k} \partial_t \chi^k = \Lambda + g_j \frac{\partial f^j}{\partial u^k} \partial_t \chi^k \\ &= -H(t, f, g) + g_j \partial_t f^j + g_j \frac{\partial f^j}{\partial u^k} \partial_t \chi^k\,. \end{aligned}$$

Differentiating the identity $q^j = f^j(t; \chi(t, q))$ with respect to t gives

$$0 = \partial_t f^j + \frac{\partial f^j}{\partial u^k} \partial_t \chi^k$$

and therefore we have $\partial_t S(t, q) = -H$. Furthermore

$$\partial_j S = \frac{\partial \sigma}{\partial u^k} \partial_j \chi^k = g_i \frac{\partial f^i}{\partial u^k} \partial_j \chi^k = g_i \delta^i_j = g_j\,,$$

which completes the proof.

At the end of this paragraph I would like to indicate the direction the work on HJ theories in mechanics has taken during the last 10 years or so: Let us go back to the Lagrangian bracket in eq. (2,17) and let us suppose again that the determinant (2,20) does not vanish. Then, for fixed t, the functions

$$q^j(u) = f^j(t; u)\,, \qquad p_j(u) = g_j(t; u)$$

define an n-dimensional submanifold of the $2n$-dimensional phase space P^{2n}. This submanifold is called a "Lagrangian" submanifold if $[u^j, u^k]_t = 0$ and will be denoted by $L_t^{(n)}$. If t changes, $L_t^{(n)}$ remains Lagrangian, because the Lagrange brackets $[u^j, u^k]$ are constants of motion, compare eq. (2,18). Lagrangian submanifolds have the following properties: For fixed t we have

$$\begin{aligned}\omega_t := \mathrm{d}p_j \wedge \mathrm{d}q^j &= \left(\frac{\partial g_j}{\partial u^k}\,\mathrm{d}u^k\right) \wedge \left(\frac{\partial f^j}{\partial u^l}\,\mathrm{d}u^l\right)\\ &= \tfrac{1}{2}[u^k, u^l]\,\mathrm{d}u^k \wedge \mathrm{d}u^l = 0 \end{aligned} \tag{2,28}$$

which shows that the concept of "Lagrangian submanifolds" is invariant under symplectic transformations. A basis of the tangent spaces $T_{(q,p)}(L_t^{(n)})$ is given by

$$l_{(j)} = \frac{\partial f^i}{\partial u^j}\,\partial_i + \frac{\partial g_i}{\partial u^j}\,\frac{\partial}{\partial p_i}$$

and eq. (2,28) means $\omega_t(l_{(j)}, l_{(k)}) = 0$ for all $j, k = 1, \ldots, n$. This shows that the tangent spaces $T_{(q,p)}(L_t^{(n)})$ are "isotropic" with respect to the symplectic form ω_t. As the dimension of $L_t^{(n)}$ is n, those tangent spaces are isotropic subspaces of maximal dimension.

Since $\omega_t = \mathrm{d}\lambda_t$, $\lambda_t = p_j\,\mathrm{d}q^j$, the relation (2,28) means that λ_t is closed on the tangent spaces of $L_t^{(n)}$. According to Poincaré's lemma we have locally

$$\lambda_t = \mathrm{d}S_t(q) = \partial_j S_t(q)\,\mathrm{d}q^j\,.$$

$S_t(q)$ is called a generating function for the Lagrangian submanifold $L_t^{(n)}$. The time evolution of $S_t(q) = S(t, q)$ is given by the HJ equation $\partial_t S(t, q) + H(t, q, p = \partial_j S) = 0$.

Relations between the n-dimensional Lagrangian submanifolds $L_t^{(n)}$ and the n-dimensional wave fronts $S(t, q) = \sigma = \text{const.}$ will be discussed in chapter 8, where we shall investigate those "singular" situations which have $\Delta(t; u) = 0, \infty$, i.e. when we have "focal" points.

2.3. Another derivation of the equations of motion

We next turn to the derivation of the canonical equations of motion by "varying" the action integral

$$A[\hat{C}] = \int_{\hat{C}} \theta\,, \qquad \theta = -H(t, q, p)\,\mathrm{d}t + p_j\,\mathrm{d}q^j\,, \qquad \hat{C}(t) = \{(t, q(t), p(t))\}\,.$$

The 1-form θ depends on $2n+1$ variables t, q^j and p_j (or v^j), $j = 1, \ldots, n$, which are the coordinates of the extended phase space $\hat{P}^{2n+1}$. Consider the 1-parameter "variation" φ_τ of $\hat{P}^{2n+1}$:

$$t \to \varphi_\tau(t), \quad q \to \varphi_\tau(q), \quad p \to \varphi_\tau(p), \quad \tau \in [0, 1],$$
$$\varphi_{\tau=0}(t) = t, \quad \varphi_{\tau=0}(q) = q, \quad \varphi_{\tau=0}(p) = p. \tag{2,29}$$

(Compare the discussion in chapter 1, preceding eq. (1,19)!) A curve $\hat{C} = (t, q(t), p(t))$ will be "deformed" by the transformation (2,29) into

$$\varphi_\tau(\hat{C}) = \{(\varphi_\tau(t), \varphi_\tau[q(\varphi_\tau(t))], \varphi_\tau[p(\varphi_\tau(t))])\}$$

and at $(\varphi_\tau(t), \varphi_\tau(q), \varphi_\tau(p))$ we have the form θ_τ, where (t, q, p) has been replaced by $(\varphi_\tau(t), \varphi_\tau(q), \varphi_\tau(p))$.

If we define

$$A_\tau := A[\varphi_\tau(\hat{C})] = \int_{\varphi_\tau(\hat{C})} \theta_\tau = \int_{\hat{C}} \varphi_\tau^* \theta,$$

then, according to eq. (1,19), we obtain

$$\left.\frac{dA_\tau}{d\tau}\right|_{\tau=0} = \int_{\hat{C}} L(V)\,\theta = \int_{\hat{C}} i(V)\,d\theta + \int_{\partial\hat{C}} i(V)\,\theta$$

where V is the vector field induced on $\hat{P}^{2n+1}$ by the "variation" (2,29):

$$V = V_{(t)}\partial_t + V^j_{(q)}\partial_j + V^{(p)}_j \frac{\partial}{\partial p_j},$$
$$V_{(t)}(t, q, p) = \partial_\tau \varphi_\tau(t)_{\tau=0}, \quad \partial_\tau := \partial/\partial\tau,$$
$$V^j_{(q)}(t, q, p) = \partial_\tau \varphi_\tau(q^j)_{\tau=0}, \tag{2,30}$$
$$V^{(p)}_j(t, q, p) = \partial_\tau \varphi_\tau(p_j)_{\tau=0}.$$

We are interested in those curves $\hat{C}_0$ for which

$$\left.\frac{dA_\tau}{d\tau}\right|_{\tau=0}[\hat{C}_0] = \int_{\hat{C}_0} i(V)\,d\theta + \int_{\partial\hat{C}_0} i(V)\,\theta = 0. \tag{2,31}$$

This important formula needs some interpretation: Consider first those variations φ_τ for which

$$\varphi_\tau(t) = t \text{ for all } \tau \text{ and } t, \quad \varphi_\tau(q(t_s)) = q(t_s) \text{ for all } \tau,$$

where t_s, $s = 1, 2$, denote the time variables of the boundary $\partial\hat{C}$. It follows that

$$V_{(t)} = 0\,, \qquad V^j_{(q)}(t_s, q(t_s), p(t_s)) = 0\,,$$

and therefore

$$\left.\frac{dA_\tau}{d\tau}\right|_{\tau=0} = \int\limits_{\hat{C}_0} i(V)\,d\theta = 0\,. \tag{2,32}$$

The last equation is supposed to hold for arbitrary smooth vector fields V, which implies that the integrand $i(V)\,d\theta$ has to vanish. This is interpreted as follows: $i(V)\,d\theta$ is a 1-form on $\hat{P}^{2n+1}$ which vanishes when applied to a tangent vector $\hat{C}'_0 = \partial_t + \dot{q}^j\partial_j + \dot{p}_j\,\partial/\partial p_j$ of the extremal $\hat{C}_0$. Thus, the implication of eq. (2.32) is:

$$\begin{aligned}
&[i(V)\,d\theta](\hat{C}'_0) = 0 \text{ for arbitrary } V\,,\\
&\qquad\qquad d\theta = -dH(t, q, p) \wedge dt + dp_j \wedge dq^j\,,\\
&\qquad\qquad \hat{C}'_0 = \partial_t + \dot{q}^j\partial_j + \dot{p}_j\,\partial/\partial p_j\,.
\end{aligned} \tag{2,33}$$

Taking for V the special vector fields ∂_j, $\partial/\partial p_j$ and ∂_t respectively, we obtain

$$[i(\partial_j)\,d\theta]\,(\hat{C}'_0) = -(\partial_j H + \dot{p}_j) = 0\,, \tag{2,34a}$$

$$[i(\partial/\partial p_j)\,d\theta]\,(\hat{C}'_0) = -\partial H/\partial p_j + \dot{q}^j = 0\,, \tag{2,34b}$$

$$[i(\partial_t)\,d\theta]\,(\hat{C}'_0) = \partial_j H\dot{q}^j + \frac{\partial H}{\partial p_j}\dot{p}^j = 0\,. \tag{2,34c}$$

The last equation obviously is a consequence of the other ones.

If we consider the coefficients of the canonical form θ as functions of v^j instead of p_j, we obtain from eq. (2,8)

$$d\theta = (-\partial_j L\,dt + dp_j) \wedge \omega^j\,. \tag{2,35}$$

Since $i(\partial_k)\,\omega^j = i(\partial_k)\,dq^j = \delta^j_k$, eq. (2,35) implies

$$[i(\partial_j)\,d\theta]\,(\hat{C}'_0) = -\partial_j L + \dot{p}_j = 0\,, \qquad p_j = \partial L/\partial v^j\,, \tag{2,36}$$

which are just the Euler–Lagrange equations. Observing that

$$dp_j = \frac{\partial p_j}{\partial v^k}\,dv^k + \cdots = \frac{\partial^2 L}{\partial v^k\,dv^j}\,dv^k + \cdots$$

we further get from eq. (2,35)

$$[i(\partial/\partial v^j)\,d\theta]\,(\hat{C}'_0) = \frac{\partial^2 L}{\partial v^j\,\partial v^k}\,(\dot{q}^k - v^k) = 0\,, \tag{2,37}$$

which implies that $\dot{q} = v$ on the extremals if the Legendre transformation is regular, i.e. if

$$\left|\left(\frac{\partial^2 L}{\partial v^i\, \partial v^k}\right)\right| \neq 0\,.$$

We next turn to the boundary term $i(V)\,\theta_{t_2} - i(V)\,\theta_{t_1}$ in eq. (2,31): Suppose the curve we consider is an extremal, then the first term $\int_{\hat{C}_0} i(V)\,\mathrm{d}\theta$ vanishes for arbitrary V and the equations

$$i(w)\,\theta_{t_s} = 0\,, \qquad s = 1, 2\,, \tag{2,38}$$

mean that the tangent vectors $w_s = w_{t_s}\partial_t + w^j_s\partial_j$ at $(t_s, q(t_s))$ are associated vectors of the form θ, the coefficient functions of which are determined by the extremals $\hat{C}_0(t)$. We saw in chapter 1 that the set of all associated vector fields of the closed form θ constitutes an n-dimensional completely integrable system the integral manifolds of which are the wave fronts $S(t, q) = \sigma$, where $S(t, q)$ is a solution of the HJ equation $\partial_t S + H(t, q, p = \partial S) = 0$.

The vector fields $w = w_t\partial_t + w^j\partial_j$ which obey the equation $i(w)\,\theta = 0$ are said to be "transversal" to the extremals "associated" with $\mathrm{d}\theta$. The reason is the following: On G^{n+1} the tangent vector $e_t = \partial_t + \dot{q}^j\partial_j$ of an extremal and the n tangent vectors $w_{(j)} = p_j\partial_t + H\partial_j$, $j = 1, \ldots, n$, of the associated wave front at a point (t, q) in general are linearly independent. This can be seen from the value of the determinant $|(e_t, w_{(1)}, \ldots, w_{(n)})|$ of the matrix

$$\begin{pmatrix} 1 & p_1 \cdots p_n \\ \dot{q}^1 & \\ \vdots & HE_n \\ \dot{q}^n & \end{pmatrix}, \tag{2,39}$$

where E_n is the unit matrix in n dimensions and where the $n + 1$ columns are the components of e_t and the $w_{(j)}$. The value of the determinant can conveniently be calculated in the following way:

Suppose

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$$

is an $(n + m)\times(n + m)$-matrix, where A_1 is an $(n \times n)$- and A_4 a nonsingular $(m \times m)$-matrix. It then follows from the relation

$$\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}\begin{pmatrix} E_n & 0 \\ -A_4^{-1}A_3 & E_m \end{pmatrix} = \begin{pmatrix} A_1 - A_2A_4^{-1}A_3 & A_2 \\ 0 & A_4 \end{pmatrix}$$

that [Satake, 1975, p. 76]

$$\left|\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}\right| = |A_4|\,|A_1 - A_2A_4^{-1}A_3|\,, \tag{2,40a}$$

or correspondingly, if A_1 is nonsingular

$$\left|\begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}\right| = |A_1|\,|A_4 - A_3A_1^{-1}A_2|\,. \tag{2,40b}$$

Applying the relation (2,40a) to the matrix (2,39) gives

$$|(e_t, w_{(1)}, \dots, w_{(n)})| = H^n(1 - q^jp_j/H) = -H^{n-1}L\,. \tag{2,41}$$

This means that the vectors e_t and $w_{(j)}$ are linearly independent where $HL \neq 0$! The singular case $HL = 0$ will be dealt with in chapter 8.

We can summarize an essential part of the above results in the following way: The 1-dimensional extremals $\hat{C}_0(t)$ are the integral manifolds associated with the $2n+1$ 1-forms $i(V)\,d\theta$ of rank $2n$ and the n-dimensional wave fronts $S(t, q) = \text{const.}$, transversal to the extremals, are the integral manifolds associated with the form θ itself!

2.4. Carathéodory's approach to the calculus of variation

As this review owes so much to Carathéodory's work on the calculus of variation, I would like to sketch some of his ideas:

In Carathédory's approach the notion of a "field" of extremals plays a central role. His essential arguments are as follows:

Suppose we have n first order differential equations

$$\dot{q}^j = \varphi^j(t, q)\,, \qquad j = 1, \dots, n\,, \tag{2,42}$$

in a region G^{n+1}, and a function $\tilde{L}(t, q, v)$ with the property that for the solutions $C_0(t)$ of the eqs. (2,42) one has $\tilde{L}[t, q(t), \varphi(t, q(t))] = 0$, but $\tilde{L}(t, q, v) > 0$ for all other near-by curves $C(t) = \{(t, q(t), v = \dot{q}(t))\}$ – the neighborhoods suitably defined –, then the curves C_0 obviously minimize the action integral

$$\tilde{A}[C] = \int_C \tilde{L}(t, q, \dot{q})\,dt\,.$$

For a given Lagrangian $L(t, q, \dot{q})$ with the action integral

$$A[C] = \int_C L(t, q, \dot{q})\,dt$$

there will be in general no functions $\varphi^j(t, q)$ with the properties just described. However, one can hope to find such "slope" functions for the "equivalent" variational problem defined by the Lagrangian

$$\begin{aligned}\tilde{L}(t, q, \dot{q}) &= L(t, q, \dot{q}) - \frac{dS}{dt}(t, q(t)) \\ &= L(t, q, \dot{q}) - \partial_t S - \dot{q}^j\partial_j S\,.\end{aligned} \tag{2,43}$$

The variational problems defined by the Lagrangians L and $\tilde{L}$ are equivalent in the following sense: One has

$$\tilde{A}[C]-\tilde{A}[C_0]=A[C]-A[C_0]\,,$$

because the value of the integral

$$\int\limits_{C}\frac{dS}{dt}\,dt=S(t_2,q(t_2))-S(t_1,q(t_1))$$

is independent of the curves $C(t)$ if the endpoints $q(t_s)$, $s=1,2$, are the same for all of them. Furthermore, we have

$$\frac{\partial^2\tilde{L}}{\partial v^j\partial v^k}=\frac{\partial^2 L}{\partial v^j\partial v^k}\,.$$

Suppose now, that there is a system of slope functions $\varphi^j(t,q)$ such that for the solutions $C_0(t)$ of the differential eqs. (2,42) the equality

$$\tilde{L}(t,q,\varphi(t,q))=L(t,q,\varphi(t,q))-\partial_t S-\varphi^j(t,q)\,\partial_j S=0 \tag{2,44}$$

holds, whereas $\tilde{L}(t,q,\dot{q})>0$ for all other curves $C(t)$ "near-by" which connect the same endpoints $q(t_1)$ and $q(t_2)$, then the curves $C_0(t)$ provide a solution of the variational problem to determine those curves which minimize $\tilde{A}[C]$. If the previous assumptions are valid, the function

$$\tilde{L}(t,q,v)=L(t,q,v)-\partial_t S-v^j\partial_j S$$

has a minimum for $v^j=\varphi^j(t,q)$. A necessary condition for this to be the case is that the partial derivatives $\partial\tilde{L}/\partial v^j$ vanish for $v^j=\varphi^j(t,q)$, which implies

$$\frac{\partial L}{\partial v^j}[t,q,\varphi(t,q)]=\partial_j S(t,q)\,. \tag{2,45}$$

The equation $\tilde{L}(t,q,v=\varphi(t,q))=0$ means

$$L[t,q,\varphi(t,q)]-\varphi^j(t,q)\frac{\partial L}{\partial v^j}[t,q,\varphi(t,q)]=\partial_t S(t,q)\,. \tag{2,46}$$

Carathéodory calls the eqs. (2,45) and (2,46) the fundamental equations of the calculus of variations [Carathéodory, 1935, §235]. Introducing canonical coordinates $p_j=(\partial L/\partial v^j)(t,q,v)$, $H=p_j v^j-L$, these equations mean

$$\psi_j(t,q)=\frac{\partial L}{\partial v^j}[t,q,v=\varphi(t,q)]=\partial_j S(t,q)\,,\qquad H[t,q,p=\psi(t,q)]=-\partial_t S(t,q)\,,$$

i.e. the function $S(t, q)$ introduced in eq. (2,43) has to be a solution of the HJ equation. A characteristic feature of Carathéodory's approach is that for him the wave fronts are at least as important as the extremals themselves: Given a solution $S(t, q)$ of the HJ equation, the solution $q(t)$ of the first order equation

$$\dot{q}^j = \frac{\partial H}{\partial p_j}[t, q, p_j = \partial_j S(t, q)] =: \varphi^j(t, q) \tag{2,47}$$

is an extremal, i.e. it is also a solution of $\dot{p}_j = -\partial_j H$. This was shown in connection with eq. (2,26).

If $C(t) = \{(t, q(t))\}$, $t \in [t_1, t_2]$ is any smooth curve inside the region G^{n+1} which is simply covered by solutions of the eqs. (2,47), i.e. by extremals, then the value of the line integral

$$\int_{t_1}^{t_2} dt \left\{ L[t, q(t), \varphi(t, q(t))] + [\dot{q}^j - \varphi^j(t, q(t))] \frac{\partial L}{\partial v^j}[t, q(t), \varphi(t, q(t))] \right\}$$

$$= \int_{t_1}^{t_2} dt \{ -H[t, q(t), \psi(t, q(t))] + \dot{q}^j \psi_j(t, q(t)) \} \tag{2,48}$$

depends only on the endpoints $q(t_1)$ and $q(t_2)$. This follows from eqs. (2,45) and (2,46) which show that the integrand in eq. (2,48) is equal to $dS(t, q(t))/dt$. The line integral (2,48) is Hilbert's famous "independent integral" [Hilbert, 1906; Gelfand and Fomin, 1963, §33; Hestenes, 1966, ch. 3].

2.5. The propagation of wave fronts

In the background of Carathéodory's canonical theory for fields is Huygens' principle (again!). This background was analyzed in an important paper by E. Hölder [1939], who used Lie's interpretation of Huygens' principle in terms of contact transformations [Lie and Scheffers, 1896, ch. 4; Lie, 1896] to illustrate the main points in the case of mechanics:

We saw that the wave fronts for a system with n degrees of freedom can be given by the equations $S(t, q) = \sigma = \text{const.}$, which means that for fixed σ the variable t becomes a function of the n coordinates q^j, $t = t(q)$, with derivatives

$$k_j = \partial_j t(q) = -\partial_j S / \partial_t S . \tag{2,49}$$

A plane tangent to the surface $S(t, q) = \sigma$ at the point (t, q) is given by

$$\tau - t = k_j(\xi^j - q^j) , \tag{2,50}$$

where $(\tau, \xi^1, \ldots, \xi^n) \in \mathbb{R}^{n+1}$ are the "running" variables of the plane. If the parameter σ changes from σ to $\hat{\sigma}$, the points (t, q) of the old surface will become points $(\hat{t}, \hat{q})$ of the new surface $S(\hat{t}, \hat{q}) = \hat{\sigma} = \text{const.}$ and the tangent plane (2,50) will go over into a tangent plane $\tau - \hat{t} = \hat{k}_j(\xi^j - \hat{q}^j)$. Such a transformation of the variables t, q^j and k_j is called a "contact" transformation. Because the relation $\hat{k}_j = \partial_j \hat{t}(\hat{q})$ has to hold on $S(\hat{t}, \hat{q}) = \hat{\sigma}$, if $k_j = \partial_j t(q)$ holds on $S(t, q) = \sigma$, a transformation $(t, q, k) \rightarrow$

$(\hat{t}, \hat{q}, \hat{k})$ is a contact transformation iff the relation

$$\mathrm{d}\hat{t} - \hat{k}_j\, \mathrm{d}\hat{q}^j = \rho(\mathrm{d}t - k_j\, \mathrm{d}q^j)\,, \qquad \rho = \rho(t, q, k)\,, \tag{2,51}$$

holds. Suppose we increase σ by $\delta\sigma$, then we have the infinitesimal transformations

$$\hat{t} = t + T(t, q, k)\,\delta\sigma\,, \tag{2,52a}$$

$$\hat{q}^j = q^j + Q^j(t, q, k)\,\delta\sigma\,, \tag{2,52b}$$

$$\hat{k}_j = k_j + K_j(t, q, k)\,\delta\sigma \tag{2,52c}$$

and the coefficient ρ in eq. (2,51) becomes $\rho = 1 + h\,\delta\sigma$. Inserting ρ and the expressions (2,52) into the eq. (2,51) and comparing the coefficients of $\delta\sigma$ gives

$$\mathrm{d}T - k_j\, \mathrm{d}Q^j - K_j\, \mathrm{d}q^j = h(\mathrm{d}t - k_j\, \mathrm{d}q^j)\,,$$

or

$$\mathrm{d}(k_j Q^j - T) = -h\, \mathrm{d}t + (hk_j - K_j)\, \mathrm{d}q^j + Q^j\, \mathrm{d}k_j\,. \tag{2,53}$$

Defining the function $F(t, q, k) = k_j Q^j - T$, the relation (2,53) implies

$$\partial_t F = -h\,, \qquad \partial_j F = hk_j - K_j\,, \qquad \partial F/\partial k_j = Q^j\,,$$

and therefore

$$T(t, q, k) = k_j\, \partial F/\partial k_j - F\,, \tag{2,54a}$$

$$Q^j(t, q, h) = \partial F/\partial k_j\,, \tag{2,54b}$$

$$K_j(t, q, k) = -\partial_j F - k_j\, \partial_t F\,, \tag{2,54c}$$

that is to say, the infinitesimal contact transformation (2,52) is generated by *one* function $F(t, q, k)$ and the σ-dependence of t, q and k is governed by the differential equations of wave motion

$$\mathrm{d}t/\mathrm{d}\sigma = k_j\, \partial F/\partial k_j - F\,, \tag{2,55a}$$

$$\mathrm{d}q^j/\mathrm{d}\sigma = \partial F/\partial k_j\,, \tag{2,55b}$$

$$\mathrm{d}k_j/\mathrm{d}\sigma = -\partial_j F - k_j\, \partial_t F\,. \tag{2,55c}$$

Suppose now we have solutions $t = t(\sigma; u)$, $q^j = q^j(\sigma; u)$ and $k_j = k_j(\sigma; u)$ of the eqs. (2,55) which depend on at least n parameters u^j, then the eqs. (2,55) imply

$$\begin{aligned} &\mathrm{d}t - k_j\,\mathrm{d}q^j = -\tilde{F}\,\mathrm{d}\sigma + G_j\,\mathrm{d}u^j\,, \\ &\tilde{F}(\sigma, u) = F(t(\sigma; u), q(\sigma; u), k(\sigma; u))\,, \\ &G_j(\sigma, u) = \partial t/\partial u^j - k_i\,\partial q^i/\partial u^j\,, \end{aligned} \tag{2,56}$$

where the coefficients $\tilde{F}$ and G_j have the property

$$\partial_\sigma G_j = -G_j\,\partial_t F\,, \tag{2,57a}$$

$$\mathrm{d}F/\mathrm{d}\sigma = \partial_\sigma \tilde{F} = -F\,\partial_t F\,. \tag{2,57b}$$

Differentiating the identity $\sigma = S(t(\sigma), q(\sigma))$ with respect to σ gives

$$1 = \partial_t S\,\mathrm{d}t/\mathrm{d}\sigma + \partial_j S\,\mathrm{d}q^j/\mathrm{d}\sigma\,,$$

or because of the eqs. (2,55a, b) and (2,49)

$$F\,\partial_t S = -1\,, \tag{2,58}$$

which combined with eq. (2,49), $k_j = -\partial_j S/\partial_t S$ replaces the HJ equation!

If we define

$$H = 1/F \tag{2,59a}$$

$$p_j = k_j/F\,, \tag{2,59b}$$

we recover the usual canonical framework of mechanics. In order to show this it is convenient to use the notation $y^j := \partial F/\partial k_j$. The functional determinant of the transformation (2,59b) has the value

$$|(\partial p_i/\partial k_j)| = |(F^{-1}(\delta^j_i - y^j k_i/F))| = -TF^{-n-1} = -TH^{n+1}\,. \tag{2,60}$$

The relation (2,56) now takes the form

$$-H\,\mathrm{d}t + p_j\,\mathrm{d}q^j = \mathrm{d}\sigma - (G_j/\tilde{F})\,\mathrm{d}u^j\,. \tag{2,61}$$

Because of the properties (2,57) we have

$$\partial_t(G_j/\tilde{F}) = \partial_\sigma(G_j/\tilde{F})\,\partial_t\sigma = (\partial_\sigma G_j/\tilde{F} - (G_j/\tilde{F}^2)\,\partial_\sigma\tilde{F})\,\partial_t\sigma = 0\,. \tag{2,62}$$

Thus, the 1-form (2,61) has the same structure as the 1-form (2,15). For $\mathrm{d}H = \mathrm{d}(1/F)$ we obtain

$$T\,\mathrm{d}H = F^{-1}(\partial_t F\,\mathrm{d}t + \partial_j F\,\mathrm{d}q^j) + y^j\,\mathrm{d}p_j\,, \tag{2,63}$$

which implies

$$T\,\partial_t H = \partial_t F/F\,, \qquad (2{,}64a)$$

$$T\,\partial_j H = \partial_j F/F\,, \qquad (2{,}64b)$$

$$T\,\partial H/\partial p_j = y^j = \partial F/\partial k_j\,. \qquad (2{,}64c)$$

As $T = k_j y^j - F$, we have

$$T(p_j\,\partial H/\partial p_j - H) = 1\,, \qquad (2{,}65)$$

i.e. $T = 1/L$. In addition we get

$$\frac{dq^j}{dt} = \frac{dq^j}{d\sigma}\Big/\frac{dt}{d\sigma} = \frac{\partial F}{\partial k_j}\Big/T = \frac{\partial H}{\partial p_j}\,,$$

$$\frac{dp_j}{dt} = \frac{dp_j}{d\sigma}\Big/\frac{dt}{d\sigma} = \frac{d(k_j/\tilde F)}{d\sigma}\Big/\frac{dt}{d\sigma} = \frac{d(k_j/\tilde F)}{d\sigma}\Big/T\,.$$

Using the relations (2,55c) and (2,57b) yields

$$\frac{d}{d\sigma}(k_j/\tilde F) = \frac{dk_j}{d\sigma}\Big/\tilde F - k_j\partial_\sigma\tilde F/\tilde F^2 = -\partial_j F/F\,,$$

and because of eq. (2,64b), we finally have

$$dp_j/dt = -\partial_j H\,.$$

Thus we see that there is a one-to-one correspondence between the canonical equations of motion for the particles and the equations of motion (2,55) for the associated wave fronts, provided we have $dt/d\sigma = 1/L \neq 0, \infty$, i.e. if t is a unique function of σ and vice versa. Whereas H generates an infinitesimal canonical transformation with group parameter t of the particles, the function $F = 1/H$ generates an infinitesimal contact transformation with group parameter σ of the wave fronts.

I find it utterly amazing that the dynamical laws of mechanics can be derived from the simple requirement that the "motion" of n-dimensional surfaces in an $(n+1)$-dimensional space should map the tangent planes through the points of an initial surface onto the corresponding tangent planes through the image points of the "moved" surface!

Example:

Suppose we have $H(\boldsymbol{x}, \boldsymbol{p}) = (1/2m)\,\boldsymbol{p}^2 + V(\boldsymbol{x})$, $\boldsymbol{x} \in \mathrm{R}^3$, then we get for $F(\boldsymbol{x}, \boldsymbol{k})$ the quadratic equation

$$F^2 V - F = -\frac{1}{2m}\boldsymbol{k}^2\,, \qquad (2{,}66)$$

with the roots

$$F_{1,2} = \frac{1}{2}\frac{1}{V} \pm \frac{1}{2|V|}(1 - 2V\boldsymbol{k}^2/m)^{1/2}\,. \qquad (2{,}67)$$

Since $(1-2Vk^2/m)=L^2/H^2\geq 0$, the root in eq. (2,67) is always real. It vanishes where $L=0$. The sign in eq. (2,67) has to be chosen such that $F=1/H$. Suppose, for instance, that $V>0$, then $H>0$ and we have

$$F_{1/2}=\frac{1}{2V}(1\pm|L|/H)=1/H,$$

or $2V=H\pm|L|$ and the choice of sign depends on the sign of L: If L is positive (weak coupling) then we have to take the minus sign, if L is negative (strong coupling) we have to take the plus sign.

We see that the points with $L=0$ separate different branches of F. Notice that the functional determinant (2,60) becomes singular, too, if $L=0$, because $T=1/L$. The deeper reasons for the criterium $L=0$ will be discussed in chapter 8.

It follows from eq. (2,57b) that $F=F_0=\text{const.}$ for the solutions of the eqs. (2,55) if $\partial_t F=0$, i.e. if F (or H) does not depend explicitly on t. Differentiating the eq. (2,66) with respect to k_j and x^j gives

$$(2FV-1)\,\partial F/\partial k_j=-k_j/m\,,\qquad (2FV-1)\,\partial_j F=-F^2\,\partial_j V$$

and the equations of wave motion (2,55) become

$$\frac{dt}{d\sigma}=-F_0/(2F_0V-1)\,,\qquad \frac{dx^j}{d\sigma}=-\frac{k_j}{m}\Big/(2F_0V-1)\,,\qquad \frac{dk_j}{d\sigma}=F_0^2\,\partial_j V/(2F_0V-1)\,.$$

Instead of solving these differential equations one can, of course, solve the corresponding canonical equations for $x^j(t)$ and $p_j(t)$ and then determine $t(\sigma)$ from – see eq. (2,11) –

$$\sigma(t)=\int^{t}\Lambda(\bar{t};u)\,d\bar{t}\,.$$

2.6. Bibliographical notes

The above remarks represent, of course, only a very small selection from the canonical framework of mechanics. Let me therefore point out some of the literature, in addition to that already mentioned above, which I have been using:

A little bit "old-fashioned", but still extremely valuable are the textbooks by Whittaker [1959] and Pars [1965]. Very influential upon the modern development of mechanics has been E. Cartan's classic "Leçons sur les invariants intégraux". Very important elements – many of which still wait for their "rediscovery" by mathematical physicists – are contained in Carathéodory's book "Calculus of variations and partial differential equations" [1935]. In Carathéodory's approach the Lagrange brackets (2,17) play an important role. A short summary of this approach was given by Boerner [1953]. Lepage's equivalences $\theta\equiv L\,dt \bmod I[\omega^j]$, $d\theta\equiv 0 \bmod I[\omega^j]$ in the context of mechanics were discussed by Boerner [1940a] and Dedecker [1951, 1957b].

Probably the best modern introduction to the canonical theory of mechanical systems is that of Arnold [1978]. The by far most extensive survey concerning modern research in mechanics is the second edition of the book by Abraham and Marsden [1978]. However, due to their heavy deployment of formal machineries its appeal to physicists may be limited! In any case, it is a very valuable guide to the modern work on dynamical systems. Chapter 5 of Abraham and Marsden contains a discussion of the ideas and the literature concerning Lagrangian submanifolds and their relation to HJ theories. Pioneering contributions to this field were made by Maslov [Maslov, 1972; Maslov and Fedoriuk, 1981] and Arnold [1967]. Other texts on this subject are [Weinstein, 1977; Guillemin and Sternberg, 1977; Kijowski and Tulczyjew, 1979]. Good modern introductions into mechanics are [Godbillon, 1969] and [Thirring, 1978]. See also [Souriau, 1970]; [Estabrook and Wahlquist, 1975] and [Hermann, 1977b].

3. Canonical theories for fields with two independent variables

We now come to the essential part of this review: the implications of generalizing Lepage's fundamental equivalence relations

$$\theta \equiv L\,\mathrm{d}t \,(\mathrm{mod}\; \mathrm{I}[\omega_j])\,, \qquad \mathrm{d}\theta \equiv 0 \,(\mathrm{mod}\; \mathrm{I}[\omega^j])\,,$$

of the canonical form $\theta = L\,\mathrm{d}t + p_j\omega^j = -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j$, as discussed in the last chapter, to field theories: As the essential new features can be seen already for systems with 2 independent variables, we first shall discuss only such field theories and we shall deal with applications to higher dimensions in chapter 6.

3.1. *The generalized Legendre transformation*

We denote by x^μ, $\mu = 1, 2$, $x = (x^1, x^2) \in \mathrm{G}^2$, the independent variables and by z^a, $a = 1, \ldots, n$, those variables of $y = (x^1, x^2, z^1, \ldots, z^n) \in \mathrm{G}^{2+n}$, which become the dependent variables $z^a = f^a(x)$ on 2-dimensional submanifolds $\Sigma^2 \subset \mathrm{G}^{2+n}$. Furthermore, the variables v^a_μ, $a = 1, \ldots, n$, $\mu = 1, 2$, become the derivatives $\partial_\mu z^a(x) = \partial_\mu f^a(x)$ on those surfaces Σ^2, especially on the extremals Σ^2_0, to be discussed below. This last property can be rephrased as follows: The Pfaffian forms $\omega^a = \mathrm{d}z^a - v^a_\mu\,\mathrm{d}x^\mu$ vanish on the extremals Σ^2_0 where $v^a_\mu = \partial_\mu z^a(x)$.

Let $L(x, z, \partial_\mu z)$ be a Lagrangian function of the action integral

$$A[\Sigma^2] = \int_{\mathrm{G}^2} L(x, z, \partial z)\,\mathrm{d}x^1\,\mathrm{d}x^2\,,$$

which is supposed to become stationary for the extremals Σ^2_0, if we vary the functions $z^a(x)$ and $\partial_\mu z^a(x)$ – this will be made more precise below –. By using the forms $\omega^a = \mathrm{d}z^a - v^a_\mu\,\mathrm{d}x^\mu$ we here can argue in the same way as in the case of mechanics:

As far as the extremals are concerned, the Lagrangian 2-form

$$\omega = L(x, z, v)\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 \tag{3,1}$$

is only one representative in an equivalence class of 2-forms, the most general of which, Ω, can be written as

$$\Omega = \omega + h^1_a\omega^a \wedge \mathrm{d}x^2 + h^2_a\,\mathrm{d}x^1 \wedge \omega^a + \tfrac{1}{2}h_{ab}\,\omega^a \wedge \omega^b$$
$$\equiv \omega(\mathrm{mod}\;\mathrm{I}[\omega^a])\,, \qquad h_{ba} = -h_{ab}\,. \tag{3,2}$$

The last term $\frac{1}{2}h_{ab}\omega^a \wedge \omega^b$ is new compared to mechanics and it is only possible, because Ω is a 2-form and if $n \geq 2$. The coefficients h^μ_a and h_{ab} in general can be arbitrary functions of x, z and v.

In mechanics the coefficients h_j in $\Omega = L\,dt + h_j\omega^j$ were determined by the requirement $d\Omega \equiv 0$ (mod $I[\omega^j]$), a property which is characteristic for HJ theories. It was Lepage's important idea to generalize this postulate to field theories: For the exterior derivative $d\Omega$ of the form (3,2) we get

$$\begin{aligned} d\Omega &= \left(\partial_a L\, dz^a + \frac{\partial L}{\partial v^a_\mu} dv^a_\mu\right) \wedge dx^1 \wedge dx^2 \\ &\quad + dh^1_a \wedge \omega^a \wedge dx^2 - h^1_a\, dv^a_1 \wedge dx^1 \wedge dx^2 + dh^2_a \wedge dx^1 \wedge \omega^a + h^2_a\, dx^1 \wedge dv^a_2 \wedge dx^2 + \tfrac{1}{2}d(h_{ab}\omega^a \wedge \omega^b) \\ &= \left(\frac{\partial L}{\partial v^a_\mu} - h^\mu_a\right) dv^a_\mu \wedge dx^1 \wedge dx^2 + 0(\text{mod } I[\omega^a])\,, \end{aligned} \tag{3,3}$$

where the equality $dz^a \wedge dx^1 \wedge dx^2 = \omega^a \wedge dx^1 \wedge dx^2$ has been used. We see that the condition $d\Omega \equiv 0$ (mod $I[\omega^a]$) is equivalent to

$$h^\mu_a = \partial L/\partial v^a_\mu =: \pi^\mu_a\,. \tag{3,4}$$

This is the same as in mechanics. New and important is, however, that the condition $d\Omega \equiv 0$ (mod $I[\omega^a]$) does not impose any restriction on the coefficients h_{ab}. This has far-reaching consequences: Inserting the relations (3,4) into the form (3,2), we get

$$\Omega = L\,dx^1 \wedge dx^2 + \pi^1_a\,\omega^a \wedge dx^2 + \pi^2_a\,dx^1 \wedge \omega^a + \tfrac{1}{2}h_{ab}\omega^a \wedge \omega^b\,. \tag{3,5}$$

Each choice of the coefficients h_{ab} defines a "canonical" form Ω_h. The implications of such a choice will become more evident when we define the canonical momenta. Before we do that, let me introduce some notational simplifications:

With

$$a^\mu := L\,dx^\mu + \pi^\mu_a\omega^a\,, \qquad \mu = 1, 2\,, \tag{3,6}$$

and $\varepsilon_{\mu\nu} = \varepsilon^{\mu\nu}$, $\varepsilon_{12} = 1$, $\varepsilon_{\nu\mu} = -\varepsilon_{\mu\nu}$ the form (3,5) can be written in the compact form

$$\begin{aligned} \Omega &= L\,dx^1 \wedge dx^2 + \pi^\mu_a\omega^a \wedge d\Sigma_\mu + \tfrac{1}{2}h_{ab}\omega^a \wedge \omega^b \\ &= a^\mu \wedge d\Sigma_\mu - L\,dx^1 \wedge dx^2 + \tfrac{1}{2}h_{ab}\omega^a \wedge \omega^b\,, \qquad d\Sigma_\mu = \varepsilon_{\mu\nu}\,dx^\nu\,. \end{aligned} \tag{3,7}$$

In chapter 2 the Legendre transformation $v^j \to p_j$, $L \to H$, was implemented by a change of basis $dt \to dt$, $\omega^j \to dq^j$ by inserting for ω^j the expression $dq^j - v^j\,dt$ and identifying the resulting negative coefficient of dt with the Hamilton-function H and the canonical momenta p_j with the corresponding coefficients of dq^j. Generalizing this procedure to the form (3,7) means: We first replace ω^a by $dz^a - v^a_\mu\,dx^\mu$ and express Ω in terms of the basis $dx^1 \wedge dx^2$, $dz^a \wedge dx^2$, $dx^1 \wedge dz^a$, $dz^a \wedge dz^b$:

$$\begin{aligned} \Omega &= [L - \pi^\mu_a v^a_\mu + \tfrac{1}{2}h_{ab}(v^a_1 v^b_2 - v^a_2 v^b_1)]\,dx^1 \wedge dx^2 + (\pi^1_a - h_{ab}v^b_2) \\ &\quad \times dz^a \wedge dx^2 + (\pi^2_a + h_{ab}v^b_1)\,dx^1 \wedge dz^a + \tfrac{1}{2}h_{ab}\,dz^a \wedge dz^b\,. \end{aligned}$$

If we define $h^{\mu\nu}_{ab} = \varepsilon^{\mu\nu} h_{ab}$, then we can rewrite the last equation as

$$\Omega = (L - \pi^{\mu}_{a} v^{a}_{\mu} + \tfrac{1}{2} h^{\mu\nu}_{ab} v^{a}_{\mu} v^{b}_{\nu})\, dx^1 \wedge dx^2 + (\pi^{\mu}_{a} - h^{\mu\nu}_{ab} v^{b}_{\nu})\, dz^a \wedge d\Sigma_{\mu} + \tfrac{1}{2} h_{ab}\, dz^a \wedge dz^b . \tag{3,8}$$

The generalized Legendre transformation, according to Lepage, defines the canonical momenta p^{μ}_{a} as the coefficients of $dz^a \wedge d\Sigma_{\mu}$:

$$p^{\mu}_{a} = \pi^{\mu}_{a} - h^{\mu\nu}_{ab} v^{b}_{\nu} \tag{3,9}$$

and the Hamilton-function H as the negative coefficient of $dx^1 \wedge dx^2$:

$$H = \pi^{\mu}_{a} v^{a}_{\mu} - \tfrac{1}{2} h^{\mu\nu}_{ab} v^{a}_{\mu} v^{b}_{\nu} - L . \tag{3,10}$$

The form Ω can now be rewritten in terms of "canonical" variables:

$$\Omega = -H\, dx^1 \wedge dx^2 + p^{\mu}_{a}\, dz^a \wedge d\Sigma_{\mu} + \tfrac{1}{2} h_{ab}\, dz^a \wedge dz^b . \tag{3,11}$$

The canonical momenta (3,9) reduce to the conventional ones, π^{μ}_{a}, if all h_{ab} vanish. In order to express the quantities H and h_{ab} as functions of the momenta p^{μ}_{a}, the Legendre transformation $v^{a}_{\mu} \to p^{\mu}_{a}$ has to be regular, i.e. the functional determinant of the Jacobi matrix $(\partial p^{\mu}_{a}/\partial v^{b}_{\nu})$ should be non-vanishing:

$$|(\partial p^{\mu}_{a}/\partial v^{b}_{\nu})| = \left| \left(\frac{\partial^2 L}{\partial v^{b}_{\nu}\, \partial v^{a}_{\mu}} - \frac{\partial h^{\mu\lambda}_{ac}}{\partial v^{b}_{\nu}} v^{c}_{\lambda} - h^{\mu\nu}_{ab} \right) \right| \neq 0 . \tag{3,12}$$

As the coefficients $h^{\mu\nu}_{ab}$ are arbitrary – up to now –, an appropriate choice can always guarantee the inequality (3,12). This freedom indicates an interesting possibility for defining a regular Legendre transformation, if the conventional one, $v^{a}_{\mu} \to \pi^{\mu}_{a}$ is singular, as, e.g., in gauge theories. We shall come back to this problem later on.

Let us assume we can solve the eqs. (3,9) for the variables v^{a}_{μ}:

$$v^{a}_{\mu} = \hat{\varphi}^{a}_{\mu}(x, z, p) .$$

Inserting these functions into $h^{\mu\nu}_{ab} = h^{\mu\nu}_{ab}(x, z, v)$, we obtain

$$\eta^{\mu\nu}_{ab}(x, z, p) := h^{\mu\nu}_{ab}(x, z, v = \hat{\varphi}(x, z, p))$$

and for the Hamilton-function H:

$$H(x, z, p) = \pi^{\mu}_{a}[x, z, v = \hat{\varphi}(x, z, p)]\, \hat{\varphi}^{a}_{\mu}(x, z, p) - \tfrac{1}{2} \eta^{\mu\nu}_{ab} \hat{\varphi}^{a}_{\mu} \hat{\varphi}^{b}_{\nu} - L[x, z, v = \hat{\varphi}(x, z, p)] .$$

From eqs. (3,10) and (3,9) we get

$$dH = v^{a}_{\mu}\, d\pi^{\mu}_{a} + \pi^{\mu}_{a}\, dv^{a}_{\mu} - \tfrac{1}{2} d(h^{\mu\nu}_{ab}\, v^{a}_{\mu}\, v^{b}_{\nu}) - dL$$

and

$$\mathrm{d}\pi^{\mu}_{a} = \mathrm{d}p^{\mu}_{a} + \mathrm{d}(h^{\mu\nu}_{ab} v^{b}_{\nu}) .$$

Combining these two equations gives

$$\mathrm{d}H = v^{a}_{\mu}\, \mathrm{d}p^{\mu}_{a} + \pi^{\mu}_{a}\, \mathrm{d}v^{a}_{\mu} + \tfrac{1}{2} v^{a}_{\mu} v^{b}_{\nu}\, \mathrm{d}h^{\mu\nu}_{ab} - \mathrm{d}L . \tag{3,13}$$

Inserting

$$\mathrm{d}H = \partial_{\mu} H\, \mathrm{d}x^{\mu} + \partial_{a} H\, \mathrm{d}z^{a} + \frac{\partial H}{\partial p^{\mu}_{a}}\, \mathrm{d}p^{\mu}_{a}$$

where

$$\partial_{\mu} H := \partial H / \partial x^{\mu} |_{z,p\ \text{fixed}}$$

and

$$\mathrm{d}L = \partial_{\mu} L\, \mathrm{d}x^{\mu} + \partial_{a} L\, \mathrm{d}z^{a} + \pi^{\mu}_{a}\, \mathrm{d}v^{a}_{\mu} ,$$

$$\mathrm{d}h^{\mu\nu}_{ab} = \mathrm{d}\eta^{\mu\nu}_{ab} = \partial_{\lambda} \eta^{\mu\nu}_{ab}\, \mathrm{d}x^{\lambda} + \partial_{c} \eta^{\mu\nu}_{ab}\, \mathrm{d}z^{c} + \frac{\partial \eta^{\mu\nu}_{ab}}{\partial p^{\lambda}_{c}}\, \mathrm{d}p^{\lambda}_{c}$$

into eq. (3,13) and comparing the coefficients of $\mathrm{d}x^{\mu}$, $\mathrm{d}z^{a}$ and $\mathrm{d}p^{\mu}_{a}$ gives

$$\partial_{\lambda} H - \tfrac{1}{2} \hat{\varphi}^{a}_{\mu}\, \hat{\varphi}^{b}_{\nu}\, \partial_{\lambda} \eta^{\mu\nu}_{ab} = -\partial_{\lambda} L , \tag{3,14a}$$

$$\partial_{c} H - \tfrac{1}{2} \hat{\varphi}^{a}_{\mu}\, \hat{\varphi}^{b}_{\nu}\, \partial_{c} \eta^{\mu\nu}_{ab} = -\partial_{c} L , \tag{3,14b}$$

$$\partial H / \partial p^{\lambda}_{c} - \tfrac{1}{2} \hat{\varphi}^{a}_{\mu} \hat{\varphi}^{b}_{\nu}\, \partial \eta^{\mu\nu}_{ab} / \partial p^{\lambda}_{c} = v^{c}_{\lambda} . \tag{3,14c}$$

3.2. The canonical field equations of motion

Up to now the variables x^{μ}, z^{a} and v^{a}_{μ} or p^{μ}_{a} have been treated as independent. This is no longer the case, if we look for the extremals $\hat{\Sigma}^{2}_{0}$, those 2-dimensional submanifolds $z^{a} = f^{a}(x)$, $p^{\mu}_{a} = g^{\mu}_{a}(x)$ for which the variational derivative $\mathrm{d}A[\hat{\Sigma}^{2}]/\mathrm{d}\tau$ – see eq. (1,19) – of the action integral

$$A[\hat{\Sigma}^{2}] = \int_{\hat{\Sigma}^{2}} \Omega$$

vanishes at $\tau = 0$:

$$\left.\frac{\mathrm{d}A_{\tau}}{\mathrm{d}\tau}\right|_{\tau=0} [\hat{\Sigma}^{2}_{0}] = \int_{\hat{\Sigma}^{2}_{0}} i(V)\, \mathrm{d}\Omega + \int_{\partial \hat{\Sigma}^{2}_{0}} i(V)\, \Omega = 0 , \tag{3,15}$$

where

$$V = V^{\mu}_{(x)}\partial_\mu + V^{a}_{(z)}\partial_a + V^{\mu}_{(p)a}\,\partial/\partial p^{\mu}_{a} \tag{3,16}$$

is an arbitrary vector field.

Arguments which run completely parallel to those in chapter 2 show that the tangent vectors

$$\hat{\Sigma}'_{(\mu)} = \partial_\mu + \partial_\mu z^a(x)\,\partial_a + \partial_\mu p^{\nu}_{a}(x)\,\partial/\partial p^{\nu}_{a} \tag{3,17}$$

of an extremal $\hat{\Sigma}^2_0$ have to obey the equations

$$[i(V)\,\mathrm{d}\Omega]\,(\hat{\Sigma}'_{(1)}, \hat{\Sigma}'_{(2)}) = 0 \tag{3,18}$$

for arbitrary vector fields V, which we take to be the independent ones ∂_μ, ∂_a and $\partial/\partial p^{\mu}_{a}$.

Since

$$\mathrm{d}\Omega = -\left(\partial_a H\,\mathrm{d}z^a + \frac{\partial H}{\partial p^{\mu}_{a}}\mathrm{d}p^{\mu}_{a}\right) \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 + \mathrm{d}p^{\mu}_{a} \wedge \mathrm{d}z^a \wedge \mathrm{d}\Sigma_\mu$$
$$+ \frac{1}{2}\left[\partial_\mu \eta_{ab}\,\mathrm{d}x^\mu + \partial_c \eta_{ab}\,\mathrm{d}z^c + \frac{\partial}{\partial p^{\mu}_{c}}\eta_{ab}\,\mathrm{d}p^{\mu}_{c}\right] \wedge \mathrm{d}z^a \wedge \mathrm{d}z^b\,, \tag{3,19}$$

we have

$$i(\partial_a)\,\mathrm{d}\Omega = -\partial_a H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 - \mathrm{d}p^{\mu}_{a} \wedge \mathrm{d}\Sigma_\mu - \partial_\mu \eta_{ab}\,\mathrm{d}x^\mu \wedge \mathrm{d}z^b + \tfrac{1}{2}\partial_a \eta_{bc}\,\mathrm{d}z^b \wedge \mathrm{d}z^c$$
$$- \partial_b \eta_{ac}\,\mathrm{d}z^b \wedge \mathrm{d}z^c - \frac{\partial \eta_{ac}}{\partial p^{\mu}_{b}}\mathrm{d}p^{\mu}_{b} \wedge \mathrm{d}z^c\,, \tag{3,20}$$

$$i(\partial/\partial p^{\mu}_{a})\,\mathrm{d}\Omega = -\frac{\partial H}{\partial p^{\mu}_{a}}\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \mathrm{d}z^a \wedge \mathrm{d}\Sigma_\mu + \frac{1}{2}\frac{\partial \eta_{bc}}{\partial p^{\mu}_{a}}\mathrm{d}z^b \wedge \mathrm{d}z^c \tag{3,21}$$

and

$$i(\partial_\mu)\,\mathrm{d}\Omega = \left(\partial_a H\,\mathrm{d}z^a + \frac{\partial H}{\partial p^{\lambda}_{a}}\mathrm{d}p^{\lambda}_{a}\right) \wedge \mathrm{d}\Sigma_\mu - \varepsilon_{\mu\nu}\,\mathrm{d}p^{\nu}_{a} \wedge \mathrm{d}z^a + \tfrac{1}{2}\partial_\mu \eta_{ab}\,\mathrm{d}z^a \wedge \mathrm{d}z^b\,. \tag{3,22}$$

Applying the 2-form (3,20) to the tangent vectors (3,17), we get the partial differential equations

$$[i(\partial_a)\,\mathrm{d}\Omega](\hat{\Sigma}'_{(1)}, \hat{\Sigma}'_{(2)}) = -\partial_a H - \partial_\mu p^{\mu}_{a} - (\partial_\mu \eta^{\mu\nu}_{ab})\partial_\nu z^b + \tfrac{1}{2}(\partial_a \eta^{\mu\nu}_{bc})\,\partial_\mu z^b\,\partial_\nu z^c$$
$$- (\partial_c \eta^{\mu\nu}_{ab})\,\partial_\mu z^c\,\partial_\nu z^b - \frac{\partial \eta^{\lambda\nu}_{ac}}{\partial p^{\mu}_{b}}\,\partial_\lambda p^{\mu}_{b}\,\partial_\nu z^c = 0\,. \tag{3,23}$$

The differential eqs. (3,23) are the analogue of the canonical equations $\dot{p}_j + \partial_j H = 0$.

If we define

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}F(x, z, p) := \partial_\mu F + \partial_a F\,\partial_\mu z^a + \frac{\partial F}{\partial p^{\nu}_{a}}\,\partial_\mu p^{\nu}_{a}\,, \tag{3,24}$$

e.g. $\mathrm{d}z^a(x)/\mathrm{d}x^\mu = \partial_\mu z^a(x)$ etc., we can rewrite the eqs. (3,23) in a more compact way:

$$\mathrm{d}p_a^\mu/\mathrm{d}x^\mu + \partial_a H = \tfrac{1}{2}\partial_a \eta_{bc}^{\mu\nu}\, \partial_\mu z^b\, \partial_\nu z^c - (\mathrm{d}\eta_{ab}^{\mu\nu}/\mathrm{d}x_\mu)\, \partial_\nu z^b\,. \tag{3,25}$$

The 2-forms (3,21) give the differential equations

$$[i(\partial/\partial p_a^\mu)\,\mathrm{d}\Omega](\hat{\Sigma}'_{(1)}, \hat{\Sigma}'_{(2)}) = -\frac{\partial H}{\partial p_a^\mu} + \partial_\mu z^a + \frac{1}{2}\frac{\partial \eta_{bc}^{\lambda\nu}}{\partial p_a^\mu}\, \partial_\lambda z^b\, \partial_\nu z^c = 0\,, \tag{3,26}$$

which are the same as the eqs. (3,14c), if $\partial_\mu z^a(x) = v_\mu^a$. We shall show below that this equality is a consequence of the inequality (3,12). The 2-forms (3,22) yield the differential equations

$$\partial_\mu z^a\, \partial_a H + \frac{\partial H}{\partial p_a^\nu}\, \partial_\mu p_a^\nu - \varepsilon_{\mu\nu}\varepsilon_{\lambda\kappa}\, \partial_\lambda p_a^\nu\, \partial_\kappa z^a + \tfrac{1}{2}(\partial_\mu \eta_{ab}^{\lambda\nu})\, \partial_\lambda z^a\, \partial_\nu z^b = 0\,. \tag{3,27}$$

These equations are, however, a consequence of the canonical eqs. (3,23) and (3,26): If we multiply eqs. (3,23) by $\partial_\mu z^a$, sum over a and use the eqs. (3,26), we obtain the eqs. (3,27).

The eqs. (3,25) and (3,26) represent a system of $n + 2n = 3n$ first-order partial differential equations for the $3n$ functions $z^a = f^a(x)$ and $p_a^\mu = g_a^\mu(x)$. By eliminating the canonical momenta p_a^μ we obtain a system of n second-order differential equations for the functions $z^a = f^a(x)$, namely the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}\frac{\partial L}{\partial v_\mu^a} - \partial_a L = 0\,. \tag{3,28}$$

This can be proven as follows: It follows from the eqs. (3,9) that

$$\frac{\mathrm{d}p_a^\mu}{\mathrm{d}x^\mu} = \frac{\mathrm{d}\pi_a^\mu}{\mathrm{d}x^\mu} - \left(\frac{\mathrm{d}}{\mathrm{d}x^\mu}\eta_{ab}^{\mu\nu}\right)\partial_\nu z^b - h_{ab}^{\mu\nu}\frac{\mathrm{d}}{\mathrm{d}x^\mu}\partial_\nu z^b\,. \tag{3,29}$$

Since $\partial_\mu\partial_\nu z^b = \partial_\nu\partial_\mu z^b$, $h_{ab}^{\mu\nu} = -h_{ab}^{\nu\mu}$, the last term on the r.h. side of eq. (3,29) vanishes. Combining the rest of the eqs. (3,29) with the canonical eqs. (3,25), we obtain

$$\mathrm{d}\pi_a^\mu/\mathrm{d}x^\mu + \partial_a H = \tfrac{1}{2}(\partial_a \eta_{bc}^{\mu\nu})\, \partial_\mu z^b\, \partial_\nu z^c\,.$$

Together with eqs. (3,14b) the last equations give the Euler–Lagrange eqs. (3,28), where the arbitrary functions h_{ab} have dropped out completely!

I next want to show that the inequality (3,12) implies $v_\mu^a = \partial_\mu z^a(x)$ on the extremals, at least in a neighborhood of $v_\mu^a = 0$: Exterior differentiation of the form (3,5) yields

$$\mathrm{d}\Omega = \partial_a L\, \mathrm{d}z^a \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 + \varepsilon_{\mu\nu}\, \mathrm{d}\pi_a^\mu \wedge \omega^a \wedge \mathrm{d}x^\nu + \tfrac{1}{2}(\mathrm{d}h_{ab}) \wedge \omega^a \wedge \omega^b + h_{ab}\, \mathrm{d}\omega^a \wedge \omega^b\,. \tag{3,30}$$

If the 2-form

$$i(\partial/\partial v_\mu^a)\,\mathrm{d}\Omega = \frac{\partial^2 L}{\partial v_\mu^a\, \partial v_\nu^b}\,\varepsilon_{\nu\lambda}\omega^b \wedge \mathrm{d}x^\lambda + \frac{1}{2}\frac{\partial h_{bc}}{\partial v_\mu^a}\,\omega^b \wedge \omega^c - h_{ab}\,\mathrm{d}x^\mu \wedge \omega^b \tag{3,31}$$

vanishes, when applied to tangent vectors $\partial_\mu + \partial_\mu z^a \partial_a + \cdots$ we obtain the equations

$$\frac{\partial^2 L}{\partial v^a_\mu\, \partial v^b_\nu} \Delta^b_\nu + \frac{1}{2} \frac{\partial h^{\lambda\nu}_{bc}}{\partial v^a_\mu} \Delta^b_\lambda \Delta^c_\nu - h^{\mu\nu}_{ab} \Delta^b_\nu = 0 \tag{3,32}$$

for the differences $\Delta^b_\nu = \partial_\nu z^b(x) - v^b_\nu$. The eqs. (3,32) constitute a set of (nonlinear) eqs. $F^\mu_a(\Delta^b_\nu) = 0$ for the quantities Δ^b_ν. The functional determinant responsible for the solvability of the eqs. (3,32) at $\Delta^b_\nu = 0$ is

$$\left| \left(\frac{\partial F^\mu_a}{\partial \Delta^b_\nu} = \frac{\partial^2 L}{\partial v^a_\mu\, \partial v^b_\nu} + \frac{\partial h^{\nu\lambda}_{bc}}{\partial v^a_\mu} \Delta^c_\lambda - h^{\mu\nu}_{ab} \right)_{\Delta^b_\nu = 0} \right| . \tag{3,33}$$

Since $\Delta^c_\lambda = -(v^c_\lambda - \partial_\lambda z^c(x))$ we have $\partial/\partial v^a_\mu = -\partial/\partial \Delta^a_\mu$ which shows that the matrix (3,33) is regular, if the determinant (3,12) does not vanish for $v^a_\mu = 0$. This implies that in a neighborhood of $\Delta^a_\mu = 0$ the eqs. (3,32) can only have the solutions $\Delta^b_\nu = 0$. ($\Delta^b_\nu = 0$ is, of course, always a solution of the eqs. (3,32), but it need not be unique.) The equality $v^a_\mu = \partial_\mu z^a(x)$ is essential for showing that the canonical eqs. (3,25) and (3,26) imply the Euler–Lagrange equations. This can be seen clearly, if we derive the Euler–Lagrange equations from the form (3,30) directly: Interior multiplication with ∂_a gives

$$i(\partial_a)\, \mathrm{d}\Omega = \partial_a L\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 - \mathrm{d}\pi^\mu_a \wedge \mathrm{d}\Sigma_\mu + h_{ba}\, \mathrm{d}\omega^b + 0(\mathrm{mod}\ \mathrm{I}[\omega^a]) . \tag{3,34}$$

With

$$\mathrm{d}\pi^\mu_a = \frac{\partial \pi^\mu_a}{\partial v^b_\nu} \mathrm{d}v^b_\nu + \partial_b \pi^\mu_a\, \mathrm{d}z^b + \partial_\nu \pi^\mu_a\, \mathrm{d}x^\nu ,$$

and because

$$\omega^a(\bar{\Sigma}'_{(\mu)}) = 0 \text{ for all } a, \mu, \text{ and } \mathrm{d}\omega^a(\bar{\Sigma}'_{(1)}, \bar{\Sigma}'_{(2)}) = 0 ,$$

if $v^a_\mu = \partial_\mu z^a$, the application of the 2-form (3,34) to the tangent vectors

$$\bar{\Sigma}'_{(\mu)} = \partial_\mu + \partial_\mu z^a \partial_a + \partial_\mu \partial_\nu z^a \frac{\partial}{\partial v^a_\nu} , \qquad \mu = 1, 2 ,$$

gives

$$[i(\partial_a)\, \mathrm{d}\Omega]\, (\bar{\Sigma}'_{(1)}, \bar{\Sigma}'_{(2)}) = \partial_a L - \frac{\partial \pi^\mu_a}{\partial v^b_\nu} \partial_\mu \partial_\nu z^b - \partial_b \pi^\mu_a\, \partial_\mu z^b - \partial_\mu \pi^\mu_a$$
$$= \partial_a L - \mathrm{d}\pi^\mu_a / \mathrm{d}x^\mu = 0 .$$

3.3. Variational and characteristic systems

Before we discuss some applications and examples, let me point out an important structural difference between the "variational" system $\{i(\partial_a)\, \mathrm{d}\Omega,\ i(\partial/\partial p^\mu_a)\, \mathrm{d}\Omega\}$ of differential forms which determine the extremals in the case of 2 independent variables as discussed in this chapter and the corresponding system $\{i(\partial_j)\,d\theta,\ i(\partial/\partial p_j)\,d\theta\}$ in mechanics, as discussed in chapter 2:

In the latter case the 1-forms $i(\partial_j)\,d\theta$, $i(\partial/\partial p_j)\,d\theta$ generate the characteristic Pfaffian system $C^*(d\theta)$ of $d\theta$, i.e. the extremals associated with the canonical form θ coincide with the characteristic integral submanifolds of the form $d\theta$, that is to say, in mechanics the tangent vectors

$$\hat{C}' = \partial_t + \dot{q}^j\partial_j + \dot{p}_j\frac{\partial}{\partial p_j} = \partial_t + \frac{\partial H}{\partial p_j}\partial_j - \frac{\partial H}{\partial q^j}\frac{\partial}{\partial p_j} \tag{3,35}$$

form a characteristic vector field associated with the form $d\theta$: The vector field (3,35) is annihilated by the forms $i(\partial_j)\,d\theta$ and $i(\partial/\partial p_j)\,d\theta$.

The situation is different for field theories: The tangent vectors $\hat{\Sigma}'_{(\mu)} = \partial_\mu + \partial_\mu z^a\partial_a + \partial_\mu p^\nu_a\,\partial/\partial p^\nu_a$ are annihilated in pairs – as a 2-vector $\hat{\Sigma}'_{(1)} \wedge \hat{\Sigma}'_{(2)}$ – by the $(3n+2)$ 2-forms

$$i(\partial_a)\,d\Omega\,, \qquad i(\partial/\partial p^\mu_a)\,d\Omega\,, \qquad i(\partial_\mu)\,d\Omega\,, \tag{3,36}$$

whereas the characteristic Pfaffian system $C^*(d\Omega)$ is generated – see the discussion following eqs. (1,23) – by

$$\begin{aligned} &i(\partial_\mu)\,i(\partial_\nu)\,d\Omega\,, \qquad i(\partial_\mu)\,i(\partial_a)\,d\Omega, \qquad i(\partial_a)\,i(\partial_b)\,d\Omega\,,\\ &i(\partial_\mu)\,i(\partial/\partial p^\nu_a)\,d\Omega\,, \qquad i(\partial_a)\,i(\partial/\partial p^\nu_b)\,d\Omega\,. \end{aligned} \tag{3,37}$$

Each characteristic vector Y of the form $d\Omega$, i.e. $i(Y)\,d\Omega = 0$, is also annihilated by the 2-forms $i(V)\,d\Omega$, where $V = \partial_\mu,\ \partial_a,\ \partial/\partial p^\mu_a$, because $i(Y)\,i(V)\,d\Omega = -i(V)\,i(Y)\,d\Omega = 0$. However, if $i(Y)\,i(V)\,d\Omega = 0$, we cannot conclude that $i(Y)\,d\Omega = 0$.

A simple example [Lepage, 1942b, p. 256] may illustrate the situation: Take $n = 1$ and let L be of the form $L = T(v) - V(z)$ which is very common in physics. We then have $p^\mu = \pi^\mu$ and

$$\begin{aligned} \Omega &= L\,dx^1 \wedge dx^2 + \pi^1\,\omega \wedge dx^2 + \pi^2\,dx^1 \wedge \omega\\ &= -H\,dx^1 \wedge dx^2 + \pi^1\,dz \wedge dx^2 + \pi^2\,dx^1 \wedge dz\,, \qquad \omega = dz - v_\mu\,dx^\mu\,, \qquad H = \pi^\mu v_\mu - L\,. \end{aligned} \tag{3,38}$$

Then

$$d\Omega = (\partial_z L\,dx^1 \wedge dx^2 - d\pi^1 \wedge dx^2 - dx^1 \wedge d\pi^2) \wedge \omega \tag{3,39}$$

is a 3-form in the 5 variables x^μ, z and v_μ, $\mu = 1, 2$. It is obvious that $\omega \in C^*(d\Omega)$. The factor

$$\begin{aligned} \rho &:= \partial_z L\,dx^1 \wedge dx^2 - d\pi^1 \wedge dx^2 - dx^1 \wedge d\pi^2\\ &= \partial_z L\,dx^1 \wedge dx^2 - \frac{\partial^2 L}{\partial v_1\,\partial v_\nu}\,dv_\nu \wedge dx^2 - \frac{\partial^2 L}{\partial v_2\,\partial v_\nu}\,dx^1 \wedge dv_\nu \end{aligned}$$

in eq. (3,39) is a 2-form in the 4 variables x^μ and v_μ. It can have the nontrivial rank 4 or 2. Since the determinant of the 4×4 skew symmetric matrix formed by the coefficients of the 6 basis vectors

$dx^1 \wedge dx^2$, $dx^\mu \wedge dv_\nu$, $dv_1 \wedge dv_2$ has the value $|(\partial^2 L/\partial v_\mu \partial v_\nu)|^2$, we see that the form ρ has the rank 4, if the matrix $(\partial^2 L/\partial v_\mu \partial v_\nu)$ is regular. In that case $d\Omega$ has the rank 5 and the integral submanifolds of the characteristic system $C^*(d\Omega)$ are 0-dimensional, whereas the integral submanifolds – the extremals – of the variational system (3,36) in our example are 2-dimensional. This shows that the situation in field theories is quite different from that in mechanics.

Let us call the system $\{i(V)\, d\Omega,\ V = \partial_\mu, \partial_a, \partial/\partial p^\mu_a\}$ of 2-forms on $G^{2+n} \times R^{2n} = \{(x^1, x^2, z^1, \ldots, z^n), (v^1_1, \ldots, v^n_2)$ or $(p^1_1, \ldots, p^2_n)\}$ the "variational" system of Ω. Only if Ω is a 1-form, its variational system coincides with the characteristic system $C^*(d\Omega)$.

The variational system generates an ideal $I[i(V)\, d\Omega]$ in the algebra of forms on $G^{2+n} \times R^{2n}$. All elements of $I[i(V)\, d\Omega]$ vanish when restricted to any extremal Σ^2_0. Since the ideal is generated by 2-forms, it will in general not be complete, that is, it will in general not contain all forms ω_Σ which vanish on the extremals Σ^2_0. For instance, the 1-forms $\omega^a = dz^a - v^a_\mu\, dx^a$ vanish on Σ^2_0, but do not belong to $I[i(V)\, d\Omega]$ [Dedecker, 1977a]. Furthermore, there are questions like this: Are the 1-dimensional integral elements $A^\mu \partial_\mu + B^a \partial_a + C^\mu_a\, \partial/\partial p^\mu_a$ regular or singular [E. Cartan, 1945, ch. IV]? Is the ideal $I[i(V)\, d\Omega]$ in involution with respect to the variables x^μ [E. Cartan, 1945, ch. V]? That these problems are not trivial can be seen from the example (i) below.

3.4. Examples and applications

The essential new element of the canonical framework introduced above are the coefficients $h_{ab}(x, z, v) = \eta_{ab}(x, z, p)$ which occur in the definition (3,9) of the canonical momenta and in the Hamilton-function (3,10). Up to now these coefficients are completely arbitrary. Before we try to identify the origin of this arbitrariness, let us discuss some examples:

(i) $h_{ab} = 0$.

In this case we have

$$p^\mu_a = \pi^\mu_a\,, \qquad H = \pi^\mu_a v^a_\mu - L\,.$$

If the Legendre transformation $v^a_\mu \to \pi^\mu_a$ is regular, i.e. if the matrix $(\partial^2 L/\partial v^a_\mu\, \partial v^b_\nu)$ is regular, we can express H as a function of x, z and π and the canonical equations (3,25) and (3,26) take the form

$$\partial_\mu \pi^\mu_a = -\partial_a H\,, \qquad \partial_\mu z^a = \frac{\partial H}{\partial \pi^\mu_a} =: \hat{\varphi}^a_\mu(x, z, \pi)\,. \tag{3,40}$$

The choice $h_{ab} = 0$ is the conventional one in physics. In the mathematical literature this case is called the canonical theory for fields of DeDonder and Weyl [DeDonder, 1913 and 1935; Weyl, 1934 and 1935]. For a system with just one real field variable this theory is the only possible one, because there can be no nonvanishing h_{ab}.

It is instructive to have a closer look – for details see [von Rieth and Kastrup, 1983] – at the variational system $I[i(V)\, d\Omega_0]$ which is generated by the 2-forms

$$i(\partial_a)\, d\Omega_0 = -\partial_a H\, dx^1 \wedge dx^2 - d\pi^\mu_a \wedge d\Sigma_\mu =: \lambda_a\,, \tag{3,41a}$$

$$i(\partial/\partial \pi^\mu_a)\, d\Omega_0 = \omega^a \wedge d\Sigma_\mu =: \omega^a_\mu\,, \qquad \omega^a = dz^a - \hat{\varphi}^a_\mu\, dx^\mu\,, \tag{3,41b}$$

$$i(\partial_\mu)\,\mathrm{d}\Omega_0 = \omega^a \wedge (\partial_a H\,\mathrm{d}\Sigma_\mu + \varepsilon_{\mu\rho}\,\mathrm{d}\pi^\rho_a) - \hat{\varphi}^a_\mu\,\lambda_a \qquad (3{,}41c)$$
$$= (\mathrm{d}H - \partial_\rho H\,\mathrm{d}x^\rho) \wedge \mathrm{d}\Sigma_\mu + \varepsilon_{\mu\rho}\,\mathrm{d}z^a \wedge \mathrm{d}\pi^\rho_a =: \omega_\mu\,.$$

The differential system (3,41) has some peculiar properties:
1. We have

$$\mathrm{d}\lambda_a = -\mathrm{d}(\partial_a H) \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2\,, \qquad \mathrm{d}\omega^a_\mu = -\mathrm{d}(\hat{\varphi}^a_\mu) \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2\,,$$
$$\mathrm{d}\omega_\mu = -\mathrm{d}(\partial_\mu H) \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2\,,$$

and since

$$\mathrm{d}x^\mu \wedge \lambda_a = \mathrm{d}\pi^\mu_a \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2\,, \qquad \mathrm{d}x^\mu \wedge \omega^a_\nu = -\delta^\mu_\nu\,\mathrm{d}z^a \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2$$

we see that $\mathrm{d}\lambda_a$, $\mathrm{d}\omega^a_\mu$ and $\mathrm{d}\omega_\mu$ belong to $\mathrm{I}[\lambda_a, \omega^a_\mu, \omega_\mu]$.
2. The $3n+2$ forms λ_a, ω^a_μ and ω_μ are linearly independent.
3. If $v = A^\mu\partial_\mu + B^a\partial_a + C^\mu_a\,\partial/\partial\pi^\mu_a$ is an arbitrary tangent vector, then it is a 1-dimensional integral element of the system (3,41) – which contains no 1-forms – and its polar system [E. Cartan, 1945, ch. IV] $i(v)\,\lambda_a$, $i(v)\,\omega^a_\mu$, $i(v)\,\omega_\mu$ has at most the rank $2n+2$! This follows from the relations

$$A^\mu\,i(v)\,\omega^a_\mu = (B^a - A^\rho\,\hat{\varphi}^a_\rho)\,A^\mu\,\mathrm{d}\Sigma_\mu$$

and

$$A^\mu\,i(v)\,\omega_\mu + B^a\,i(v)\,\lambda_a + C^\mu_a\,i(v)\,\omega^a_\mu = i(v)\,i(v)\,\mathrm{d}\Omega_0 = 0\,,$$

which show that at most $2n+2$ of the 1-forms $i(v)\,\lambda_a$, $i(v)\,\omega^a_\mu$, $i(v)\,\omega_\mu$ are linearly independent. The maximal rank $2n+2$ is realized, e.g., for the vector $v = \partial/\partial x^1$.
4. For $v_{(\mu)} = \partial_\mu + \hat{\varphi}^a_\mu\,\partial_a + C^\nu_{\mu,a}\,\partial/\partial\pi^\nu_a$ the rank of its polar system is at most $2n+1$, because now $\delta^\mu_\nu\,i(v_{(\nu)})\,\omega^a_\mu = 0$. Thus the vectors $v_{(\mu)}$ are singular integral elements [E. Cartan, 1945, p. 64].
5. The rank of the reduced polar system [E. Cartan, 1945, ch. V], which is obtained from the polar system by setting $\mathrm{d}x^\mu = 0$, $\mu = 1, 2$, in $i(v)\,\lambda_a$, $i(v)\omega^a_\mu$ and $i(v)\,\omega_\mu$, is at most $2n+1$ and $2n$ for the $v_{(\mu)}$ of the previous number 4. Thus the variational system (3,41) is *not* in involution with respect to the variables x^1 and x^2 [E. Cartan, 1945, ch. V]!
6. The following "prolongation" avoids the above "diseases": If one adjoins the 1-forms ω^a – which we know to vanish on the extremals – and their exterior derivatives $\mathrm{d}\omega^a$ to the ideal $\mathrm{I}[\lambda_a, \omega^a_\mu, \omega_\mu]$, then the new ideal is generated by the $3n$ forms ω^a, $\mathrm{d}\omega^a$ and λ_a, because ω^a_μ and ω_μ lie in the ideal $\mathrm{I}[\omega^a, \mathrm{d}\omega^a, \lambda_a]$. The rank s_0 of the linear equations $i(v)\,\omega^a = 0$ in general will be n and the solutions v are of the form

$$v = A^\mu\,\partial_\mu + A^\rho\,\hat{\varphi}^a_\rho\,\partial_a + C^\rho_a\,\partial/\partial\pi^\rho_a\,, \qquad (3{,}42)$$

where the $2n+2$ coefficients A^μ, C^μ_a are arbitrary.

The maximal rank $s_0 + s_1$ of the polar system ω^a, $i(v)\,\mathrm{d}\omega^a$, $i(v)\,\lambda_a$ with v as in eq. (3,42) is* $3n$. As the

*provided the matrix $(\partial^2 H/\partial\pi^\mu_a\,\partial\pi^\nu_b)$ fulfills certain regularity conditions.

maximal rank of the corresponding reduced polar system is $3n$, too, the differential system generated by the forms ω^a, $d\omega^a$ and λ_a is in involution with respect to the variables x! Let us, therefore, call the differential system $I[\omega^a, d\omega^a, \lambda_a]$ the "proper variational system"!

7. We remark further that $d\Omega_0 = \omega^a \wedge \lambda_a$ which implies $\omega^a \wedge d\lambda_a = d\omega^a \wedge \lambda_a$.

(ii) $h_{ab} = \text{const.} \neq 0$.

In this case we have $p^\mu_a \neq \pi^\mu_a$, but the terms in the equations of motion (3,25) and (3,26) involving the parameters h_{ab} explicitly drop out. The reason is the following: For $h_{ab} = \text{const.}$ the last term in eq. (3,11) can be written as

$$\tfrac{1}{2}h_{ab}\, dz^a \wedge dz^b = \tfrac{1}{2}d(h_{ab}z^a\, dz^b)\,, \tag{3,43}$$

and therefore this term does not contribute to $d\Omega$.

In more conventional language the terms involving the parameters h_{ab} may be interpreted as follows [Debever, 1941]: Defining

$$\frac{d}{dx^\mu} := \partial_\mu + v^a_\mu\, \partial_a + v^a_{\mu\nu}\frac{\partial}{\partial v^a_\nu}\,, \qquad v^a_{\nu\mu} = v^a_{\mu\nu}\,,$$

we have, with $h^{\mu\nu}_{ab} = \varepsilon^{\mu\nu}h_{ab}$,

$$h^{\mu\nu}_{ab}\, v^a_\mu\, v^b_\nu = h^{\mu\nu}_{ab}\, d(z^a v^b_\nu)/dx^\mu\,.$$

If in addition $dh^{\mu\nu}_{ab}/dx^\mu = 0$, e.g. if the $h^{\mu\nu}_{ab}$ are constants, then

$$h^{\mu\nu}_{ab} v^a_\mu v^b_\nu = \frac{d}{dx^\mu}(h^{\mu\nu}_{ab} z^a v^b_\nu)\,, \tag{3,44}$$

i.e. the term $h^{\mu\nu}_{ab} v^a_\mu v^b_\nu$ is a total divergence! This implies that the Lagrangian

$$L^*(x, z, v) = L(x, z, v) - \tfrac{1}{2}h^{\mu\nu}_{ab}\, v^a_\mu\, v^b_\nu \tag{3,45}$$

gives the same field equations (3,28) as L itself, but the canonical momenta are

$$\partial L^*/\partial v^a_\mu = \pi^\mu_a - h^{\mu\nu}_{ab}\, v^b_\nu = p^\mu_a\,.$$

The procedure of adding a "surface" term (3,44) to a given Lagrangian L can be used in gauge theories in order to construct a regular Legendre transformation $v^a_\mu \to p^\mu_a$, when $v^a_\mu \to \pi^\mu_a$ is singular. Example: E-dynamics in 1 space and 1 time dimension: For this system we have

$$\begin{aligned} &L = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - j_\alpha A^\alpha = \tfrac{1}{2}(F_{01})^2 - j_\alpha A^\alpha\,,\\ &x = (x^0, x^1)\,, \qquad x\cdot x = (x^0)^2 - (x^1)^2\,,\\ &F_{01} = -F_{10} = \partial_0 A_1 - \partial_1 A_0 = E\,,\\ &v^\alpha_\mu = \partial_\mu A^\alpha\,, \qquad \mu, \alpha = 0, 1\,, \end{aligned} \tag{3,46}$$

from which we obtain

$$\pi_0^0 = 0\,, \qquad \pi_0^1 = \partial_0 A^1 + \partial_1 A^0 = -E\,,$$
$$\pi_1^0 = \partial_0 A^1 + \partial_1 A^0 = -E\,, \qquad \pi_1^1 = 0\,.$$

For the coefficients $h_{\alpha\beta}$ we choose

$$h_{\alpha\beta} = -\lambda \varepsilon_{\alpha\beta}\,, \qquad \lambda = \text{const}\,.$$

We then get for the canonical momenta (3,9):

$$p_0^0 = \lambda\, \partial_1 A^1\,, \qquad p_1^0 = (1-\lambda)\,\partial_1 A^0 + \partial_0 A^1\,,$$
$$p_0^1 = (1-\lambda)\,\partial_0 A^1 + \partial_1 A^0\,, \qquad p_1^1 = \lambda\, \partial_0 A^0\,. \tag{3,47}$$

The determinant of this Legendre transformation has the value $\lambda^3(2-\lambda)$ [Dedecker, 1977], i.e. the transformation is singular for $\lambda = 0$ – which is well-known – and for $\lambda = 2$ – a case which will be discussed in chapter 5 –. Let us discuss the choice $\lambda = 1$ in more detail. We here have

$$p_0^0 = \partial_1 A^1\,, \qquad p_1^0 = \partial_0 A^1\,,$$
$$p_0^1 = \partial_1 A^0\,, \qquad p_1^1 = \partial_0 A^0 \tag{3,48}$$

and

$$H = \tfrac{1}{2}(p_1^0 + p_0^1)^2 + (p_0^0 p_1^1 - p_1^0 p_0^1) + j_0 A^\alpha$$
$$= \tfrac{1}{2}((p_1^0)^2 + (p_0^1)^2) + p_0^0 p_1^1 + j_\alpha A^\alpha\,. \tag{3,49}$$

The canonical equations $\partial_\mu A^\alpha = \partial H/\partial p_\alpha^\mu$ are just the eqs. (3,48) and the equations

$$\partial_\mu p_\alpha^\mu = -\partial H/\partial A^\alpha = -j_\alpha$$

are the Maxwell equations

$$\partial_1 E = j^0\,, \qquad \partial_0 E = -j^1\,.$$

The Lagrangian (3,45) takes the form

$$L^* = L + \varepsilon_{\alpha\beta}\,\partial_0 A^\alpha\,\partial_1 A^\beta$$
$$= -\tfrac{1}{2}\partial_\mu A^\alpha\,\partial^\mu A_\alpha + \tfrac{1}{2}(\partial_\mu A^\mu)^2 - j_\alpha A^\alpha\,. \tag{3,50}$$

Of special interest in the context of this simple gauge theory is the behavior of the form

$$\Omega = L\,\mathrm{d}x^0 \wedge \mathrm{d}x^1 + \pi_\alpha^0\,\omega^\alpha \wedge \mathrm{d}x^1 + \pi_\alpha^1\,\mathrm{d}x^0 \wedge \omega^\alpha + \tfrac{1}{2}h_{\alpha\beta}\,\omega^\alpha \wedge \omega^\beta \tag{3,51}$$

under gauge transformations $A^\alpha \to A^\alpha + \partial^\alpha f(x)$:

If the variables v^{α}_{μ} transform as

$$v^{\alpha}_{\mu} \to v^{\alpha}_{\mu} + \partial^{\alpha}\partial_{\mu} f(x),$$

then the forms $\omega^{\alpha} = \mathrm{d}A^{\alpha} - v^{\alpha}_{\mu}\,\mathrm{d}x^{\mu}$ are gauge invariant! As the momenta π^{μ}_{α} are gauge invariant, too, the form (3,51) is gauge invariant, if the Lagrangian L and the coefficients $h_{\alpha\beta}$ are gauge invariant. Thus, for $j_{\alpha} = 0$ and our choice $h_{\alpha\beta} = -\varepsilon_{\alpha\beta}$ the form (3,51) is gauge invariant!

On the other hand, in the canonical representation,

$$\Omega = -H\,\mathrm{d}x^0 \wedge \mathrm{d}x^1 + p^0_{\alpha}\,\mathrm{d}A^{\alpha} \wedge \mathrm{d}x^1 + p^1_{\alpha}\,\mathrm{d}x^0 \wedge \mathrm{d}A^{\alpha} + \tfrac{1}{2}h_{\alpha\beta}\,\mathrm{d}A^{\alpha} \wedge \mathrm{d}A^{\beta} \tag{3,52}$$

each individual term is gauge dependent, whereas their sum is gauge invariant, because it equals the sum in eqs. (3,51)!

This behavior of the form Ω under gauge transformations becomes even more striking, if we consider a system, where the potentials A^{α} are coupled to a complex scalar matter field φ with the gauge invariant Lagrangian

$$L = -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} + (\partial_{\mu} - \mathrm{i}qA_{\mu})\varphi^* \cdot (\partial^{\mu} + \mathrm{i}qA^{\mu})\varphi - \mu^2\varphi^*\varphi\,, \tag{3,53}$$

for which we obtain

$$\begin{aligned} \pi^{\mu} &= \partial L/\partial\partial_{\mu}\varphi = (\partial^{\mu} - \mathrm{i}qA^{\mu})\,\varphi^*\,, \\ \bar{\pi}^{\mu} &= \partial L/\partial\partial_{\mu}\varphi^* = (\partial^{\mu} + \mathrm{i}qA^{\mu})\varphi\,. \end{aligned} \tag{3,54}$$

We define the forms

$$\omega_{\varphi} = \mathrm{d}\varphi - \varphi_{\mu}\,\mathrm{d}x^{\mu}\,, \qquad \bar{\omega}_{\varphi} = \mathrm{d}\varphi^* - \bar{\varphi}_{\mu}\,\mathrm{d}x^{\mu}\,,$$

and assume that the quantities φ_{μ}, $\bar{\varphi}_{\mu}$ transform under gauge transformations

$$A^{\alpha} \to A^{\alpha} + \partial^{\alpha} f(x)\,, \qquad \varphi \to \exp\{-\mathrm{i}q\,f(x)\}\,\varphi\,, \qquad \varphi^* \to \exp\{\mathrm{i}q\,f(x)\}\,\varphi^*$$

as follows

$$\begin{aligned} \varphi_{\mu} &\to \exp\{-\mathrm{i}q\,f(x)\}\,(\varphi_{\mu} - \mathrm{i}q\,\partial_{\mu}f(x)\,\varphi)\,, \\ \bar{\varphi}_{\mu} &\to \exp\{\mathrm{i}q\,f(x)\}\,(\bar{\varphi}_{\mu} + \mathrm{i}q\,\partial_{\mu}f(x)\,\varphi^*)\,. \end{aligned} \tag{3,55}$$

Then the forms ω_{φ}, $\bar{\omega}_{\varphi}$ transform "covariantly":

$$\omega_{\varphi} \to \exp\{-\mathrm{i}q\,f(x)\}\,\omega_{\varphi}\,, \qquad \bar{\omega}_{\varphi} \to \exp\{\mathrm{i}q\,f(x)\}\,\bar{\omega}_{\varphi}$$

and since

$$\pi^{\mu} \to \exp\{\mathrm{i}q\,f(x)\}\,\pi^{\mu}\,, \qquad \bar{\pi}^{\mu} \to \exp\{-\mathrm{i}q\,f(x)\}\,\bar{\pi}^{\mu}\,,$$

the terms $\pi^\mu \omega_\varphi$ and $\bar{\pi}^\mu \bar{\omega}_\varphi$ are invariant under gauge transformations! Thus, the form

$$\Omega = L\,\mathrm{d}x^0 \wedge \mathrm{d}x^1 + \pi^\mu_\alpha\, \omega^\alpha \wedge \mathrm{d}\Sigma_\mu - \tfrac{1}{2}\lambda \varepsilon_{\alpha\beta}\, \omega^\alpha \wedge \omega^\beta + \pi^\mu\, \omega_\varphi \wedge \mathrm{d}\Sigma_\mu + \bar{\pi}^\mu\, \bar{\omega}_\varphi \wedge \mathrm{d}\Sigma_\mu \tag{3,56}$$

is gauge invariant!

Let me conclude this preliminary discussion of applying the general canonical framework discussed above to 2-dimensional E-dynamics by the following remark:
Since the field $F_{\alpha\beta}$ is antisymmetric in the indices α and β, it is possible to use $h_{\alpha\beta} = \lambda F_{\alpha\beta}$, $\lambda = \text{const.}$! Since $F_{01} = -(\partial_0 A^1 + \partial_1 A^0)$, this is an example for h_{ab} depending on v^a_μ. For the canonical momenta (3,9) we get in this case

$$\begin{aligned} p^0_0 &= \lambda(\partial_1 A^0 + \partial_0 A^1)\,\partial_1 A^1\,, \qquad & p^0_1 &= (\partial_1 A^0 + \partial_0 A^1)(1 - \lambda\,\partial_1 A^0)\,,\\ p^1_0 &= (\partial_1 A^0 + \partial_0 A^1)(1 - \lambda\,\partial_0 A^1)\,, \qquad & p^1_1 &= \lambda(\partial_1 A^0 + \partial_0 A^1)\,\partial_0 A^0\,, \end{aligned} \tag{3,57}$$

i.e. the Legendre transformation is nonlinear. Its functional determinant has the (gauge invariant) value

$$2\lambda^3(\partial_0 A^1 + \partial_1 A^0)^3\,(1 - \lambda(\partial_0 A^1 + \partial_1 A^0)) = -2\lambda^3 E^3(1 + \lambda E)\,.$$

It follows from the eqs. (3,57) that

$$p^0_1 + p^1_0 = 2(\partial_1 A^0 + \partial_0 A^1) - \lambda(\partial_1 A^0 + \partial_0 A^1)^2\,.$$

The roots of this equation are

$$\partial_1 A^0 + \partial_0 A^1 = \frac{1}{\lambda}(1 \pm [1 - \lambda(p^0_1 + p^1_0)]^{1/2})\,. \tag{3,58}$$

Inserting the values (3,58) for $\partial_1 A^0 + \partial_0 A^1$ into the eqs. (3,57) immediately gives the quantities $\partial_\mu A^\alpha$ as functions of the momenta p^μ_α which allows to calculate $H = H(x, A, p)$ etc.

(iii) A nontrivial unique choice of the coefficients h_{ab} is obtained as follows: By means of the 1-forms

$$a^\mu = L\,\mathrm{d}x^\mu + \pi^\mu_a \omega^a\,, \qquad \theta^\mu = -H\,\mathrm{d}x^\mu + p^\mu_a\,\mathrm{d}z^a \tag{3,59}$$

we can construct the following form of type (3,5) or (3,8) respectively:

$$\begin{aligned} \Omega_c &= \frac{1}{L}\, a^1 \wedge a^2\\ &= L\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \pi^\mu_a\, \omega^a \wedge \mathrm{d}\Sigma_\mu + \frac{1}{2L}(\pi^1_a \pi^2_b - \pi^2_a \pi^1_b)\, \omega^a \wedge \omega^b\\ &= -\frac{1}{H}\,\theta^1 \wedge \theta^2 \qquad\qquad (3,60)\\ &= -H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + p^\mu_a\,\mathrm{d}z^a \wedge \mathrm{d}\Sigma^\mu - \frac{1}{2H}(p^1_a p^2_b - p^2_a p^1_b)\,\mathrm{d}z^a \wedge \mathrm{d}z^b\,. \end{aligned}$$

The form Ω_c is unique among all possible forms (3,5), because it is the only one which has the minimal rank 2! That the forms (3,60) have rank 2 is obvious. On the other hand, assume that the form (3,5) has rank 2. Then there exist 2 Pfaffian forms $\rho^{(\mu)}$, $\mu = 1, 2$ such that $\Omega = \rho^{(1)} \wedge \rho^{(2)}$, which implies $\Omega \wedge \Omega = 0$. Writing out this last equation for the expression (3,5), or (3,8) respectively, the coefficients of $dx^1 \wedge dx^2 \wedge \omega^a \wedge \omega^b$, or $dx^1 \wedge dx^2 \wedge dz^a \wedge dz^b$ respectively, have to vanish. This gives the same values for h_{ab} as in eqs. (3,60).

The form (3,60) defines Carathéodory's canonical theory for fields, which will be discussed in detail in chapter 5.

3.5. The classification of canonical theories according to the rank of their basic canonical 2-form

Let me make some preliminary comments on the origin of the possibility for introducing those arbitrary coefficients h_{ab} into the canonical reformulation of a system defined by a Lagrangian L: First clues one can see already in mechanics: If we replace a given Lagrangian $L(t, q, \dot{q})$ by $L^*(t, q, \dot{q}) = L(t, q, \dot{q}) + df(q)/dt = L(t, q, \dot{q}) + \dot{q}^j \partial_j f$, then the Euler–Lagrange equations of L and L^* are the same, i.e. the second-order differential equations, which determine the dynamics of the system, are not affected by the additional term df/dt in L. However, for the canonical momenta p_j we have

$$p_j^* = \frac{\partial L^*}{\partial \dot{q}_j} = \partial L/\partial \dot{q}_j + \partial_j f = p_j + \partial_j f ,$$

that is to say, the term df/dt does change the definition of the canonical momenta! We see therefore that already in mechanics the canonical reformulation of n second-order differential equations by a system of $2n$ first-order differential equations is not unique. We have seen that in field theories the corresponding freedom is much more substantial if $n \geq 2$!

In view of this large freedom to introduce different canonical theories for fields, depending on the choice of the coefficients h_{ab}, the question arises, how to classify these theories. Probably the most important classifying criterion is the rank of the form Ω for a given set of coefficients h_{ab}. The reason is the following: Since $d\Omega = 0$ on the extremals Σ_0^2, the rank r of Ω coincides with its class, that is to say, the integral submanifolds associated with the form Ω itself have the dimension $n + 2 - r$, $r \geq 2$. In mechanics the integral submanifolds associated with the canonical form θ are the n-dimensional wave fronts $S(t, q) = \text{constant}$. We see that the rank of the canonical form Ω is decisive for the structure and dimension of the "wave fronts" associated with a given "family" of extremals!

Before we discuss the general case let us look at some examples:

(i) If $n = 1$, i.e. if we have only one dependent variable z, the rank of Ω is always 2, because the rank of a p-form over a $(p + 1)$-dimensional space is always p [Godbillon, 1969, p. 29].

(ii) If all h_{ab} vanish, the DeDonder–Weyl canonical form

$$\Omega_0 = a^\mu \wedge d\Sigma_\mu - L\, dx^1 \wedge dx^2$$

– see eq. (3,7) – has rank 4 (for $n \geq 2$), because it is expressed by the 4 linearly independent Pfaffian forms dx^μ and a^μ. Thus, the integral submanifolds of Ω_0 in general have the dimension $2 + n - 4 = n - 2$!

(iii) Carathéodory's canonical form (3,60) has rank 2. It is the only canonical form which allows for n-dimensional wave fronts!

(iv) Consider the case $n = 2$, h_{12} arbitrary. Then the determinant of the skew symmetric coefficient matrix

$$\begin{pmatrix} 0 & -H & p_2^1 & p_2^2 \\ H & 0 & -p_1^1 & -p_1^2 \\ -p_2^1 & p_1^1 & 0 & \eta_{12} \\ -p_2^2 & p_1^2 & -\eta_{12} & 0 \end{pmatrix}$$

of the canonical form

$$\Omega = -H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + p_a^\mu\,\mathrm{d}z^a \wedge \mathrm{d}\Sigma_\mu + \eta_{12}\,\mathrm{d}z^1 \wedge \mathrm{d}z^2 \tag{3,61}$$

has the value

$$(H\eta_{12} + |p|)^2, \qquad |p| := p_1^1 p_2^2 - p_2^1 p_1^2 .$$

Thus, the form (3,61) has rank 4 iff $H\eta_{12} + |p| \neq 0$ (the singular case $\eta_{12} = -|p|/H$ is just that of Carathéodory, compare eq. (3,60)). The associated wave fronts are therefore 0-dimensional!

Let us turn to the general case now. For this purpose we express the forms Ω in terms of the canonical variables p, η and H by means of the forms $\theta^\mu = -H\,\mathrm{d}x^\mu + p_a^\mu\,\mathrm{d}z^a$:

$$\Omega = \theta^\mu \wedge \mathrm{d}\Sigma_\mu + H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \tfrac{1}{2}\eta_{ab}\,\mathrm{d}z^a \wedge \mathrm{d}z^b . \tag{3,62}$$

Let $w = w^{(\mu)}\partial_\mu + w^a\partial_a$ be a vector field associated with the form Ω. This means that w has to obey the equation

$$i(w)\,\Omega = -\theta^2(w)\,\mathrm{d}x^1 + \theta^1(w)\,\mathrm{d}x^2 + (w^{(1)}p_a^2 - w^{(2)}p_a^1 + w^b\eta_{ba})\,\mathrm{d}z^a = 0$$

which implies the $n+2$ homogeneous equations

$$\theta^\mu(w) = -Hw^{(\mu)} + p_a^\mu w^a = 0\,, \qquad \mu = 1, 2\,, \tag{3,63a}$$

$$w^{(1)}p_a^2 - w^{(2)}p_a^1 + w^b\eta_{ba} = 0\,, \qquad a = 1, \dots, n \tag{3,63b}$$

for the components $w^{(\mu)}$, w^a.

We know from chapter 1 that there are $n+2-r$ linearly independent vector fields $w_{(\sigma)}$, $\sigma = 1, \dots, n+2-r$, if Ω has the rank r. Thus the matrix of the eqs. (3,63) must have the rank r, too.

In the case of Carathéodory's form (3,60) the associated vector fields have to obey only the 2 eqs. (3,63a). A system of n linearly independent solutions of these equations are

$$w_{(a)} = p_a^\mu\,\partial_\mu + H\,\partial_a\,, \qquad a = 1, \dots, n\,. \tag{3,64}$$

We shall see in chapter 5 that in Carathéodory's canonical theory at each point $(x, z) \in G^{2+n}$ the tangent vectors (3,64) of a wave front and the tangent vectors $e_{(\mu)} = (\delta_\mu^1, \delta_\mu^2, v_\mu^1, \dots, v_\mu^n)$ of an extremal span an $(n+2)$-dimensional vector space, if $HL \neq 0$, because the determinant

$$|(e_{(1)}, e_{(2)}, w_{(1)}, \ldots, w_{(n)})| = \left| \begin{pmatrix} E_2 & p \\ v & HE_n \end{pmatrix} \right| , \qquad p = (p^\mu_a) , \qquad v = (v^a_\mu)$$

has the value $-LH_c^{n-1}$, in complete analogy to eq. (2.41) from mechanics. Thus, Carathéodory's canonical theory for fields has the same transversality property as the canonical theory of mechanics. This is no longer the case, if the rank of Ω is larger than 2! In the DeDonder–Weyl case, where all h_{ab} vanish, the eqs. (3,63b) imply $w^{(\mu)} = 0$, which means that the corresponding wave fronts lie in the planes x^μ = const. The other components w^a have to obey the two equations $w^a p^\mu_a = 0$. Thus, there can be only $n-2$ linearly independent vector fields associated with the form Ω_0 and we loose the nice transversality properties we have in mechanics and in Carathéodory's theory!

The rank of the form Ω will be essential for the properties of Hamilton–Jacobi theories for fields to be discussed in the next chapter.

3.6. Bibliographical notes

Terms of the type $h_{ab}(v^a_1 v^b_2 - v^a_2 v^b_1)$ where first discussed in 1859 by Clebsch in a paper on the "2. variation" of a variational integral with several independent variables. Clebsch noticed that for constant h_{ab} the integrand $h_{ab}(v^a_1 v^b_2 - v^a_2 v^b_1)$ in the action integral could be transformed "away" into a surface term, which nevertheless would change the matrix $(\partial^2 L/\partial v^a_\mu \partial v^a_\nu)$, which is essential for any discussion of the 2. variation, into $(\partial^2 L/\partial v^a_\mu \partial v^b_\nu + h^{\mu\nu}_{ab})$.

More than 40 years later the subject was taken up by Hadamard [1902 and 1905], again in the context of the 2. variation.

In 1929 Carathéodory's fundamental paper on a canonical theory for fields appeared, in which the coefficients h_{ab} have the value $(\pi^1_a \pi^2_b - \pi^2_a \pi^1_b)/L$. This theory was further developed by Boerner [1936, 1940a, b, 1953] and in an important contribution by E. Hölder [1939].

Lepage was the first, who, in a series of fundamental papers in the midthirties and early forties, discussed the general case in terms of Cartan's exterior differential calculus [1936a, b, 1941 and 1942a, b]. Whereas Lepage analyzed the subject merely on the local level, it is mainly the merit of Dedecker, to investigate the global aspects (bundle structure, algebraic-topological properties etc.) of Lepage's ideas [1950, 1951, 1952a and b, 1953, 1957a and b, 1977a–d, 1978]. An interesting contribution to global aspects of Lepage's theory is due to Liesen [1967].

Exterior differential $(p+1)$-forms Ω_{p+1} and the integral submanifolds of "associated" p-forms $i(Y)\,\Omega_{p+1}$ were discussed by Gallissot in the context of fluid mechanics [1958]. See also Nôno and Mimura [1972 and 1975].

As to the DeDonder–Weyl calculus of variations for fields see Klötzler [1970]. An exposition of the DW theory in terms of modern differential geometry was given by Goldschmidt and Sternberg [1973].

In a series of interesting papers a "Warsaw" school of mathematical physicists has developed and investigated a "multi-symplectic" or "multi-phase-space" approach to classical field theories in the framework of modern differential geometry. It uses – essentially – the DeDonder–Weyl form Ω_0 as a multi-symplectic canonical form, with some modifications in mind [Trautmann, 1967, 1972; Sniatycki, 1970; Kijowski, 1972, 1973, 1974, 1977; Kijowski and Szczyrba, 1976; Kijowski and Tulczyjew, 1979; Gawedzki, 1972; Gawedzki and Kondracki, 1974; Tulczyjew, 1974; Szczyrba, 1974, 1976, 1977, 1981]. Additional references can be found in [Kijowski and Tulczyjew, 1979].

A related approach is due to García and Pérez-Rendón [García and Pérez-Rendón, 1969, 1971, 1978; García, 1974, 1977]. See also García et al. (eds.) [1980] for additional papers on the subject. The

DeDonder–Weyl canonical form Ω_0 has been discussed by Aldaya and de Azcárraga [1980] too. Recent discussions on global aspects of Lagrangian field theories and their Hamiltonian structures are contained in papers by Gelfand and Dikii [1975, 1977], Takens [1977, 1979], Vinogradov [1977, 1978], Kupershmidt [1980] and Anderson and Duchamp [1980].

4. Hamilton–Jacobi theories for fields

4.1. The HJ theory for fields of DeDonder and Weyl

The main purpose Lepage had in mind when introducing the equivalence relation $\mathrm{d}\Omega \equiv 0$ (mod $\mathrm{I}[\omega^a]$) was the concept of Hamilton–Jacobi theories for fields: We saw in mechanics that the property $\mathrm{d}\theta = 0$, $\theta = -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j$, on the extremals, combined with Poincaré's lemma (see ch. 1), implies the basic HJ relation $\mathrm{d}S(t, q) = -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j$. In the same way we can conclude from $\mathrm{d}\Omega \equiv 0$ (mod $\mathrm{I}[\omega^a]$) that Ω is locally an exact 2-form on the extremals Σ_0^2, that is to say Ω is expressable by differentials $\mathrm{d}S^1(x, z)$, $\mathrm{d}S^2(x, z)$, $\mathrm{d}x^1$, $\mathrm{d}x^2, \ldots$, where the number of these differentials necessary for expressing Ω is equal to the rank of Ω. Since Ω is a closed form, its rank is equal to its class.

For instance, in the case of Carathéodory's theory – see eqs. (3,60) –, where $\Omega = \Omega_c$ has rank 2, we have

$$\mathrm{d}S^1(x, z) \wedge \mathrm{d}S^2(x, z) = -\frac{1}{H_c}\,\theta^1 \wedge \theta^2\,. \tag{4,1}$$

We shall discuss this CHJ relation in detail in the next chapter.

As the DeDonder–Weyl form $\Omega = \Omega_0$ has rank 4, we here have

$$\begin{aligned}\Omega_0 &= \mathrm{d}S^1(x, z) \wedge \mathrm{d}x^2 + \mathrm{d}x^1 \wedge \mathrm{d}S^2(x, z)\\ &= -H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \pi_a^\mu\,\mathrm{d}z^a \wedge \mathrm{d}\Sigma_\mu\end{aligned} \tag{4,2}$$

which implies

$$\pi_a^\mu = \partial_a S^\mu(x, z) =: \psi_a^\mu(x, z)\,, \tag{4,3a}$$

$$\partial_\mu S^\mu(x, z) = H(x, z, \pi = \psi(x, z))\,. \tag{4,3b}$$

Eqs. (4,3) are obviously simple generalizations of the relations $p_j = \partial_j S$, $\partial_t S + H = 0$ in mechanics. We discussed already in the last chapter – and we can infer it again from the expression (4,2) – that the form Ω has deficiencies with respect to the transversality properties between extremals and the wave fronts to be calculated from the eqs. (4,3): According to eq. (4,2) the wave fronts are given by $S^\mu(x, z) = \sigma^\mu =$ const., *and* $x^\mu =$ const., because

$$i(w)\,\Omega_0 = \mathrm{d}S^1(w)\,\mathrm{d}x^2 - w^{(2)}\,\mathrm{d}S^1 + w^{(1)}\,\mathrm{d}S^2 - \mathrm{d}S^2(w)\,\mathrm{d}x^1 = 0\,,$$

$$w = w^{(\mu)}\partial_{(\mu)} + w^a\partial_a\,, \qquad \partial_{(\mu)} := \partial/\partial x^\mu\,,$$

$$\mathrm{d}S^\mu(w) = w^{(\nu)}\,\partial_{(\nu)}S^\mu + w^a\,\partial_a S^\mu$$

implies

$$w^{(\mu)} = 0\,, \qquad \mu = 1,2\,, \qquad w^a\,\partial_a S^\mu = w^a \pi^\mu_a = 0\,,$$

which means that the wave fronts associated with Ω_0 are $(n-2)$-dimensional, not n-dimensional and that they lie in the "characteristic" planes x^μ = constant.

Despite these transversality problems the eqs. (4,3) have a rather simple form which allows one to illustrate a number of important aspects of HJ theories for fields in a rather straightforward way:

Eq. (4,3b) is a first-order partial differential equation for *two* functions $S^\mu(x, z)$. Thus, we can choose one of these functions, say $S^2(x, z)$, with a large degree of arbitrariness and then solve eq. (4,3b) for $S^1(x, z)$. The main restriction to be imposed on any choice of S^2 is the "transversality condition" that at each point $(x, z = f(x))$ the derivatives $\partial_a S^2(x, z)$ have to be equal to the canonical momenta $\pi^\mu_a(x)$ of the extremals under investigation.

There is another important difference between HJ theories in mechanics and those in field theories which should be stressed from the very beginning: We saw in chapter 2 that *any* solution of the HJ equation leads to a system

$$\dot{q}^j = \varphi^j(t, q) = \frac{\partial H}{\partial p_j}(t, q, p = \partial S(t, q))$$

of first-order differential equations, the solutions of which, $q^j(t) = f^j(t; u)$, $u = (u^1, \ldots, u^n)$, constitute an n-parametric family of extremals which generate $S(t, q)$, if $|(\partial q^j/\partial u^k)| \neq 0$: We can calculate $S(t, q)$ by computing

$$\sigma(t, u) = \int^t \mathrm{d}\bar{t}\, L(\bar{t}, f(\bar{t}; u), \partial_t f(\bar{t}; u))\,,$$

solving the equations $q^j = f^j(t; u)$ for u^k: $u^k = \chi^k(t, q)$ and inserting the functions $\chi^k(t, q)$ into $\sigma(t, u)$: $S(t, q) = \sigma(t, \chi(t, q))$.

This procedure is in general no longer possible in the case of field theories. The reason is the following: In the DeDonder–Weyl theory we have $v^a_\mu = \partial H/\partial \pi^\mu_a$. Given now any solution $S^\mu(x, z)$ of the DWHJ eq. (4,3b), we can define the "slope" functions

$$\varphi^a_\mu(x, z) := \frac{\partial H}{\partial \pi^\mu_a}(x, z, \pi^\mu_a = \partial_a S^\mu(x, z))\,. \tag{4,4}$$

However, the partial differential equations

$$\partial_\mu z^a(x) = \varphi^a_\mu(x, z) \tag{4,5}$$

will only have solutions $z^a = f^a(x)$, if the integrability conditions

$$\frac{\mathrm{d}}{\mathrm{d}x^\nu}\varphi^a_\mu(x, z(x)) = \partial_\nu \varphi^a_\mu + \partial_b \varphi^a_\mu \cdot \varphi^b_\nu = \frac{\mathrm{d}}{\mathrm{d}x^\mu}\varphi^a_\nu(x, z(x)) = \partial_\mu \varphi^a_\nu + \partial_b \varphi^a_\nu \cdot \varphi^b_\mu \tag{4,6}$$

are fulfilled! The conditions (4,6) impose stringent restrictions on the solutions $S^\mu(x, z)$, which can already be seen from the following simple example:

From $L = \frac{1}{2}(v_0)^2 - \frac{1}{2}(v_1)^2 - V(z)$ we obtain $H = \frac{1}{2}(p^0)^2 - \frac{1}{2}(p^1)^2 + V(z)$ and the DWHJ equation

$$\partial_\mu S^\mu + \tfrac{1}{2}(\partial_z S^0)^2 - \tfrac{1}{2}(\partial_z S^1)^2 + V(z) = 0\,. \tag{4,7}$$

Furthermore, we have

$$\varphi_0(x, z) = \partial_z S^0(x, z)\,, \qquad \varphi_1(x, z) = -\partial_z S^1(x, z)\,, \tag{4,8}$$

and the integrability conditions (4,6) mean here

$$-\partial_z^2 S^0\, \partial_z S^1 + \partial_1 \partial_z S^0 = -\partial_z^2 S^1\, \partial_z S^0 - \partial_0 \partial_z S^1\,. \tag{4,9}$$

The "separating" ansatz

$$S^\mu(x, z) = h^\mu(x) + W^\mu(z) \tag{4,10}$$

gives for the eq. (4,7):

$$\partial_\mu h^\mu(x) + \tfrac{1}{2}(\partial_z W^0)^2 - \tfrac{1}{2}(\partial_z W^1)^2 + V(z) = 0 \tag{4,11}$$

and for the conditions (4,9):

$$W^{0\prime\prime} W^{1\prime} = W^{1\prime\prime} W^{0\prime}\,, \qquad W^{0\prime} := \partial_z W^0 \text{ etc}\,. \tag{4,12}$$

This equation has the solution

$$W^0 = a W^1 + b\,, \qquad a, b = \text{const.}\,, \tag{4,13}$$

where the additive constant b is irrelevant and can be taken as $b = 0$. Because $v_0 = W^{0\prime}$ $v_1 = -W^1$, the ansatz (4,10) allows only for extremals $z(x)$ with the property $\partial_0 z(x)/\partial_1 z(x) = \text{const.}$, i.e. the extremals are of the plane-wave type $z(x) = f(\alpha x^0 + \beta x^1)$, α, β = const.

It follows from the eq. (4,11) that

$$-\partial_\mu h^\mu(x) = \tfrac{1}{2}A = \text{const.} = \tfrac{1}{2}(\partial_z W^0)^2 - \tfrac{1}{2}(\partial_z W^1)^2 + V(z)\,. \tag{4,14}$$

This equation and eq. (4,12) can be satisfied by the ansatz

$$S^0 = -\tfrac{1}{4}Ax^0 + \omega\, W(z)\,, \qquad S^1 = -\tfrac{1}{4}Ax^1 + k\, W(z)\,, \tag{4,15}$$

where ω, k = const. With $\mu^2 = \omega^2 - k^2$, the eq. (4,14) implies

$$W'(z) = \frac{1}{\mu}(A - 2V(z))^{1/2}\,, \qquad W(z) = \frac{1}{\mu}\int^z \mathrm{d}\bar{z}\,(A - 2V(\bar{z}))^{1/2}\,. \tag{4,16}$$

Well-known examples for $V(z)$ are:

$$V(z) = \tfrac{1}{2}\mu^2 z^2 , \tag{4,17a}$$

$$V(z) = \tfrac{1}{2}\lambda (z^2 - a^2)^2 , \tag{4,17b}$$

$$V(z) = \alpha(1 - \cos(\beta z)) , \tag{4,17c}$$

the associated field equations of which have, of course, not only solutions of the type $z = f(\omega x^0 - kx^1)$.

Without the restrictions (4,9) it is easy to obtain more general solutions of eq. (4,11). For instance, we can take for $W^1(z)$ an arbitrary function and then determine $W^0(z)$ from

$$W^{0\prime}(z) = [A + (W^{1\prime})^2 - 2V(z)]^{1/2} .$$

Thus, there are wave fronts which cannot be generated by extremals!

Suppose now that we have found solutions $S^\mu(x, z)$ of the DWHJ eq. (4,3b) such that the slope functions (4,4) do obey the integrability conditions (4,6), then the solutions $z^a = f^a(x)$ of the first-order eqs. (4,5) are extremals, i.e. they are solutions of the equations.

$$\mathrm{d}\pi^\mu_a/\mathrm{d}x^\mu = -\partial_a H , \tag{4,18}$$

too.

Proof: Since

$$\pi^\mu_a(x) = \psi^\mu_a(x, z(x)) = \partial_a S^\mu(x, z(x)) ,$$

we get

$$\mathrm{d}\pi^\mu_a/\mathrm{d}x^\mu = \partial_a\partial_\mu S^\mu(x, z(x)) + \partial_b\partial_a S^\mu(x, z(x))\, \partial_\mu z^b(x) .$$

On the other hand we have

$$\partial_\mu S^\mu = -H[x, z(x), \pi = \partial S(x, z(x))]$$

and

$$\frac{\mathrm{d}H}{\mathrm{d}z^a} \equiv \mathrm{D}_a H := \partial_a H + \frac{\partial H}{\partial \pi^\nu_b} \partial_b\partial_a S^\nu . \tag{4,19}$$

We therefore get

$$\frac{\mathrm{d}\pi^\mu_a}{\mathrm{d}x^\mu} = -\partial_a H + \left(\partial_\mu z^b(x) - \frac{\partial H}{\partial \pi^\mu_b}\right) \partial_a\partial_b S^\mu \tag{4,20}$$

and see that the canonical eqs. (4,18) are a consequence of the eqs. (4,4).

In the context of the integrability conditions (4,6) the following notion appears to be useful: We shall call an extremal $z^a = f^a(x)$ "weakly embedded" in a "geodesic field" $S^\mu(x, z)$, i.e. in a solution of the DWHJ eq. (4,3b), if the equality (4,3a) holds for the points $(x, z = f(x))$. Such a geodesic field does not have to obey the eqs. (4,6) outside $z = f(x)$. If, however, the functions $S^\mu(x, z)$ do fulfill the conditions (4,6) in an open neighborhood of $(x, f(x))$, we shall call the extremals, or a family of them, "strongly embedded" in the wave fronts $S^\mu(x, z)$. It can be shown that any extremal in general can be embedded weakly (see below), but not strongly.

4.2. Conserved current associated with a parameter-dependent solution of the DWHJ equation

There are other important properties of the HJ-theory in mechanics which allow for generalizations in field theories:

If $S(t, q; a)$ is a solution of the HJ equation which depends on a parameter a, then the quantity

$$G(t, q; a) := \partial S(t, q; a)/\partial a \tag{4,21}$$

is a constant (of motion) along such an extremal for which the relation $p_j(t) = \partial_j S(t, q(t); a)$ holds [Whittaker, 1959, §148; the statement of Gelfand and Fomin, 1963, p. 90, that the quantity G in eq. (4,21) is a constant of motion for *any* extremal is not correct !]:

$$\frac{\mathrm{d}}{\mathrm{d}t} G(t, q(t); a) = 0\,, \tag{4,22}$$

where $q^j(t) = f^j(t)$ is an extremal. In the language of differential forms the proof of this important theorem goes as follows: We have

$$\begin{aligned} \mathrm{d}S(t, q; a) &= \partial_t S\,\mathrm{d}t + \partial_j S\,\mathrm{d}q^j + (\partial S/\partial a)\,\mathrm{d}a \\ &= -H(t, q, p = \psi(t, q))\,\mathrm{d}t + \psi_j(t, q; a)\,\mathrm{d}q^j + G\,\mathrm{d}a\,, \qquad \psi_j = \partial_j S(t, q; a)\,. \end{aligned} \tag{4,23}$$

Exterior differentiation of eq. (4,23) gives

$$0 = -\left[\partial_j H\,\mathrm{d}q^j + \frac{\partial H}{\partial p_j}\mathrm{d}\psi_j(t, q; a)\right] \wedge \mathrm{d}t + \mathrm{d}\psi_j(t, q; a) \wedge \mathrm{d}q^j + \mathrm{d}G \wedge \mathrm{d}a\,. \tag{4,24}$$

Suppose $q^j = q^j(t)$ is an arbitrary curve, not necessarily an extremal, then, because of

$$\mathrm{d}\psi_j = \partial_t\psi_j\,\mathrm{d}t + \partial_k\psi_j\,\mathrm{d}q^k + \frac{\partial\psi_j}{\partial a}\mathrm{d}a\,, \qquad \mathrm{d}q^j = \dot{q}^j\,\mathrm{d}t\,,$$

it follows from eq. (4,24) that

$$\left[\dot{q}^j - \frac{\partial H}{\partial p_j}(t, q, \psi(t, q))\right]\frac{\partial\psi_j}{\partial a} - \frac{\mathrm{d}}{\mathrm{d}t}G = 0\,. \tag{4,25}$$

Thus, if the curve $q(t)$ obeys the equations $\dot{q}^j = \partial H/\partial p_j$, eq. (4,22) follows immediately from eq. (4,25).

Noether's theorem is a special case of the important property (4,22): Let

$$t \to \hat{t}(t, q; \alpha), \qquad q^j \to \hat{q}^j = \hat{q}^j(t, q; \alpha)$$

be a 1-parameter transformation group of $G^{1+n} = \{(t, q)\}$ into itself such that

$$\hat{t}(t, q; \alpha = 0) = t, \qquad \hat{q}^j(t, q; \alpha = 0) = q^j$$

and $dS(\hat{t}, \hat{q}) = dS(t, q)$. Then $\hat{S}(t, q; \alpha) = S(\hat{t}, \hat{q})$ is a solution of the HJ equation, if $S(t, q)$ is a solution.

We can identify the group parameter α in $S(t; q; \alpha)$ with the parameter a of the above theorem and obtain

$$G = \frac{\partial \hat{S}}{\partial \alpha} = \frac{\partial S}{\partial \hat{t}} \frac{\partial \hat{t}}{\partial \alpha} + \frac{\partial S}{\partial \hat{q}^j} \frac{\partial \hat{q}^j}{\partial \alpha}.$$

Defining

$$T(t, q) = \left. \frac{\partial \hat{t}}{\partial \alpha} \right|_{\alpha=0}, \qquad Q^j(t, q) = \left. \frac{\partial \hat{q}^j}{\partial \alpha} \right|_{\alpha=0}, \tag{4,26}$$

and observing that

$$\left. \frac{\partial S}{\partial \hat{t}} \right|_{\alpha=0} = \partial_t S, \qquad \left. \frac{\partial S}{\partial \hat{q}^j} \right|_{\alpha=0} = \partial_j S,$$

we obtain for G at $\alpha = 0$

$$G = \partial_t S T + \partial_j S Q^j = -HT + p_j Q^j, \tag{4,27}$$

which is just the usual expression obtained by the standard derivations of Noether's theorem [e.g. Gelfand and Fomin, 1963, section 20].

The conservation of the quantity (4,27) can be derived in a more compact way by observing that the invariance of $dS(t, q)$ under the 1-parameter transformation group is equivalent to the statement that the form $dS(t, q)$ is invariant with respect to the vector field $Y = T\partial_t + Q^j\partial_j$, which means – see the discussion following eq. (1,16) – that

$$L(Y)\, dS = 0.$$

Combining this equation with the formula (1,14) gives

$$\begin{aligned} d i(Y)\, dS &= d(T\,\partial_t S + Q^j\,\partial_j S) \\ &= d(-HT + p_j Q^j) = 0, \end{aligned} \tag{4,28}$$

which is the required result.

In order to make the preceding discussion more complete, I briefly recall the derivation of Noether's theorem without reference to the HJ theory: suppose that

$$L(Y)\,\theta = \mathrm{d}h\,, \qquad \theta = -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j\,,$$

where $h = h(t, q, p)$ may be a function of t, q and p, depending on the (Killing) vector Y (in the discussion above we made the simplifying assumption that $\mathrm{d}h = 0$). Because – see eq. (1,14) – $L(Y)\,\theta = i(Y)\,\mathrm{d}\theta + \mathrm{d}i(Y)\,\theta$ and since $i(Y)\,\mathrm{d}\theta = 0$ on the extremals – see eq. (2,33) – we have

$$\mathrm{d}[i(Y)\,\theta - h] = \mathrm{d}(-TH + Q^j p_j - h) = 0\,,$$

i.e. $-HT + p_jQ^j - h$ is a constant of motion!

The generalization of eq. (4,22) to field theories takes the following form:

If a solution S^μ of the DWHJ eq. (4,3b) depends on a parameter a, $S^\mu = S^\mu(x, z; a)$, then the functions

$$G^\mu(x, z; a) = \partial S^\mu/\partial a\,, \qquad \mu = 1, 2\,, \tag{4,29}$$

are the components of a current which is conserved "along" the extremals $z^b = f^b(x)$:

$$\frac{\mathrm{d}G^\mu}{\mathrm{d}x^\mu}(x, z = f(x); a) = 0\,. \tag{4,30}$$

The proof is completely similar to that in mechanics:

Inserting $\mathrm{d}S^\mu = \partial_\nu S^\mu\,\mathrm{d}x^\nu + \partial_b S^\mu\,\mathrm{d}z^b + (\partial S^\mu/\partial a)\,\mathrm{d}a$ into $\mathrm{d}S^1 \wedge \mathrm{d}x^2 + \mathrm{d}x^1 \wedge \mathrm{d}S^2$ yields

$$\begin{aligned}\mathrm{d}S^1 \wedge \mathrm{d}x^2 + \mathrm{d}x^1 \wedge \mathrm{d}S^2 &= \partial_\mu S^\mu\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \partial_b S^\mu\,\mathrm{d}z^b \wedge \mathrm{d}\Sigma_\mu + G^\mu\,\mathrm{d}a \wedge \mathrm{d}\Sigma_\mu \\ &= -H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \psi^\mu_b\,\mathrm{d}z^b \wedge \mathrm{d}\Sigma_\mu - G^\mu\,\mathrm{d}\Sigma_\mu \wedge \mathrm{d}a\,, \qquad \mathrm{d}\Sigma_\mu = \varepsilon_{\mu\nu}\,\mathrm{d}x^\nu\,.\end{aligned} \tag{4,31}$$

Exterior differentiation of eq. (4,31) gives

$$0 = -\left(\partial_b H\,\mathrm{d}z^b + \frac{\partial H}{\partial \pi^\mu_b}\,\mathrm{d}\psi^\mu_b(x, z; a)\right) \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2 + \mathrm{d}\psi^\mu_b \wedge \mathrm{d}z^b \wedge \mathrm{d}\Sigma_\mu - \mathrm{d}G^\mu \wedge \mathrm{d}\Sigma_\mu \wedge \mathrm{d}a\,. \tag{4,32}$$

If $z^b = f^b(x)$ is an arbitrary smooth function of x, then eq. (4,32) becomes

$$0 = \left[\left(\partial_\mu z^b(x) - \frac{\partial H}{\partial \pi^\mu_b}\right)\frac{\partial \psi^\mu_b}{\partial a} - \frac{\mathrm{d}G^\mu}{\mathrm{d}x^\mu}\right]\mathrm{d}a \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2\,. \tag{4,33}$$

Since an extremal $z^b = f^b(x)$ obeys the equations $\partial_\mu z^b - \partial H/\partial \pi^\mu_b = 0$, the continuity eq. (4,30) follows immediately from eq. (4,33).

Examples:

The solution (4,15) depends on the parameters A and k. The current associated with the parameter A is

$$G^0_A = \frac{\partial S^0}{\partial a} = -\frac{1}{4}x^0 + \frac{1}{2}\frac{\omega}{\mu}\int \mathrm{d}z(A - 2V(z))^{-1/2}\,,$$
$$G^1_A = \frac{\partial S^1}{\partial a} = -\frac{1}{4}x^1 + \frac{1}{2}\frac{k}{\omega}\int \mathrm{d}z\,(A - 2V(z))^{-1/2}\,. \tag{4,34}$$

Since

$$\frac{\mathrm{d}G^0_A}{\mathrm{d}x^0} = -\frac{1}{4} + \frac{1}{2}\frac{\omega}{\mu}(A - 2V(z))^{-1/2}\,\partial_0 z(x) = -\frac{1}{4} + \frac{1}{2}\frac{\omega^2}{\mu^2}\,,$$

$$\frac{\mathrm{d}G^1_A}{\mathrm{d}x^1} = -\frac{1}{4} - \frac{1}{2}\frac{k^2}{\mu^2}\,,$$

the current (4,34) is indeed conserved. For $V(z) = \frac{1}{2}\mu^2 z^2$, $A > 0$ and the extremals $z = (\sqrt{A}/\mu)\sin(\omega x^0 - kx^1)$, the current (4,34) becomes trivial:

$$G^0_A = -\frac{1}{4}x^0 + \frac{\omega}{2\mu^2}\arcsin(\mu z/\sqrt{A}) = -\frac{1}{4}x^0 + \frac{\omega}{2\mu^2}(\omega x^0 - kx^1)\,,$$

$$G^1_A = -\frac{1}{4}x^1 + \frac{k}{2\mu^2}(\omega x^0 - kx^1)\,.$$

The conserved current associated with the parameter k is

$$G^0_k = \frac{\partial S^0}{\partial k} = \frac{k}{\omega}W(z)\,, \qquad G^1_k = \frac{\partial S^1}{\partial k} = W(z)\,. \tag{4,35}$$

It is important to keep in mind that the currents (4,29) in general are not conserved for arbitrary solutions of the field equations, but only for those which are embedded at least weakly in the special geodesic field $S^\mu(x, z; a)$: For the first term in eq. (4,33) to vanish it is necessary and sufficient that we have $\partial_\mu f^b = \varphi^b_\mu(x, z)$ for $z^b = f^b(x)$, where $\varphi^b_\mu(x, z)$ is given by eq. (4,4). The above examples of conserved currents may not look very interesting. However, there is the following intriguing possibility: Imagine that one can find a solution $S^\mu(x, z)$ which depends on an arbitrary function. Then it is possible, in principle at least, to generate an infinite number of conserved currents. This might be of interest for the integration of the system (Takhtadzhyan and Fadeev, 1974; Pohlmeyer, 1976; Thacker, 1981].

Before we come to Noether's theorem for field theories, let me make two remarks:

(i) In terms of differential forms the continuity equation for a current G^μ can be formulated in the following way: If we define the current 1-form

$$G = G^\mu\,\mathrm{d}\Sigma_\mu\,, \tag{4,36}$$

the continuity equation $\mathrm{d}G^\mu/\mathrm{d}x^\mu = 0$ can be written as

$$\mathrm{d}G = \frac{\mathrm{d}G^\mu}{\mathrm{d}x^\mu}\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 = 0\,. \tag{4,37}$$

(ii) Suppose $z^a = f^a(x)$ is a function such that the Legendre transformed $\partial_\mu z^a(x)$, namely $\pi^\mu_a(x) = \partial_a S^\mu(x, z = f(x))$, obeys the eqs. (4,3), then we have the relation

$$L(x, z(x), \partial_\mu z(x)) = \pi^\mu_a \partial_\mu z^a - H = \partial_a S^\mu \partial_\mu z^a + \partial_\mu S^\mu = \frac{\mathrm{d}S^\mu}{\mathrm{d}x^\mu}(x, z(x)) . \tag{4,38}$$

Historically the eq. (4,38) was the starting point for the Hamilton–Jacobi theory for fields introduced by DeDonder and Weyl, respectively: It allows one to express the integral $\int L\, \mathrm{d}x^1\, \mathrm{d}x^2$ over a region G by the integral $\int S^\mu\, \mathrm{d}\Sigma_\mu$ over the boundary ∂G!

Noether's theorem for fields can be obtained as follows: Let

$$x^\mu \to \hat{x}^\mu = \hat{x}^\mu(x, z; \alpha), \qquad z^a \to \hat{z}^a = \hat{z}^a(x, z; \alpha),$$
$$\hat{x}^\mu(x, z; 0) = x^\mu, \qquad \hat{z}^a(x, z; 0) = z^a,$$

be a 1-parameter transformation group which leaves

$$\tilde{\Omega}_0 = \mathrm{d}S^1 \wedge \mathrm{d}x^2 + \mathrm{d}x^1 \wedge \mathrm{d}S^2$$

invariant. If we define the Killing vector

$$X^\mu = \left.\frac{\partial \hat{x}^\mu}{\partial \alpha}\right|_{\alpha=0}, \qquad Z^a = \left.\frac{\partial \hat{z}^a}{\partial \alpha}\right|_{\alpha=0}, \qquad Y = X^\mu \partial_\mu + Z^a \partial_a$$

the invariance of $\tilde{\Omega}_0$ means

$$L(Y)\, \tilde{\Omega}_0 = 0 ,$$

or, because of eq. (1,14)

$$\mathrm{d}(i(Y)\, \tilde{\Omega}_0) = 0 , \tag{4,39}$$

which implies that $i(Y)\, \tilde{\Omega}_0$ is a conserved current 1-form on the extremals:

For $z^a = f^a(x)$ we have

$$\begin{aligned} i(Y)\, \tilde{\Omega}_0 &= G^\mu\, \mathrm{d}\Sigma_\mu , \\ G^\mu &= \mathrm{d}S^\mu(Y) + \varepsilon^{\mu\nu} \varepsilon_{\lambda\rho} X^\lambda\, \mathrm{d}S^\rho/\mathrm{d}x^\nu \\ &= X^\nu (\partial_\nu S^\mu + \delta^\mu_\nu\, \mathrm{d}S^\rho/\mathrm{d}x^\rho - \mathrm{d}S^\mu/\mathrm{d}x^\nu) + Z^a\, \partial_a S^\mu . \end{aligned} \tag{4,40}$$

Using the relation (4,38) and

$$T^\mu_\nu = \pi^\mu_a v^a_\nu - \delta^\mu_\nu L = \mathrm{d}S^\mu/\mathrm{d}x^\nu - \partial_\nu S^\mu - \delta^\mu_\nu\, \mathrm{d}S^\rho/\mathrm{d}x^\rho , \tag{4,41}$$

the components G^μ of the current (4,40) can be written as

$$G^\mu = -T^\mu_\nu X^\nu + \pi^\mu_a Z^a . \tag{4,42}$$

Without using the HJ theory Noether's theorem is obtained as follows: Suppose that

$$L(Y)\,\Omega_0 = \mathrm{d}h\,, \qquad \Omega_0 = a^\mu\,\mathrm{d}\Sigma_\mu - L\,\mathrm{d}x^1 \wedge \mathrm{d}x^2\,, \qquad a^\mu = L\,\mathrm{d}x^\mu + \pi^\mu_a \omega^a\,,$$

where $h = h^\mu(x, z, \pi)\,\mathrm{d}\Sigma_\mu$ is a 1-form depending on the Killing vector Y. As $L(Y)\,\Omega_0 = \mathrm{d}i(Y)\,\Omega_0 = \mathrm{d}h$ on the extremals, where $i(Y)\,\mathrm{d}\Omega_0 = 0$, we have for the extremals $z^a = f^a(x)$ the current conservation

$$\mathrm{d}G = 0\,, \qquad G = i(Y)\,\Omega_0 - h = G^\mu\,\mathrm{d}\Sigma_\mu$$
$$= (-X^\rho T^\mu_\rho + Z^a \pi^\mu_a - h^\mu)\,\mathrm{d}\Sigma_\mu\,.$$

4.3. The "complete" integral of the DWHJ equation

We next come to the important concept of a "complete" HJ integral for fields. Let me recall the corresponding integral in mechanics: If $S(t, q; a)$ is a solution of the HJ equation which depends on n parameters a_j, $j = 1, \ldots, n$, such that

$$\left| \left(\frac{\partial^2 S}{\partial q^j \partial a_k} = \frac{\partial \psi_j}{\partial a_k}(t, q; a) \right) \right| \neq 0\,, \tag{4,43}$$

then we can solve the equations

$$\partial S/\partial a_k = b^k = \text{const.}\,, \qquad k = 1, \ldots, n\,,$$

for the coordinates q^j: $q^j(t) = f^j(t; a, b)$. The curves $q(t) = f(t; a, b)$ are extremals. This follows from eq. (4,25), where $G = b^k = \text{const.}$ now and which implies

$$(\dot{q}^j - \partial H/\partial p_j)\,\partial\psi_j/\partial a_k = 0\,, \qquad k = 1, \ldots, n\,. \tag{4,44}$$

In view of the inequality (4,43) the coefficients $\dot{q}^j - \partial H/\partial p_j$ of the homogeneous system (4,44) have to vanish, which means that the functions $q^j(t) = f^j(t; a, b)$ are solutions of the canonical equations

$$\dot{q}^j = \frac{\partial H}{\partial p_j}(t, q, p_j = \partial_j S(t, q))\,.$$

It follows from our discussion in chapter 2 that they are solutions of $\dot{p}_j = -\partial_j H$, too! A solution $S(t, q; a)$ of the HJ equation with the property (4,43) is called a "complete integral". It provides a set of solutions $q^j(t) = f^j(t; a, b)$ of the equations of motion which depends on the largest possible number ($2n$) of constants of integration, i.e. the set is "complete".

The definition of a "complete" integral $S^\mu(x, z)$ of HJ equations in field theories is a straightforward generalization of that in mechanics [Fréchet, 1905; DeDonder, 1935; Dedecker, 1953], however, it will in general not provide a "complete" set of solutions of the field equations.

Suppose, we have a solution $S^\mu(x, z)$ of eq. (4,3b) which depends on $2n$ parameters a^ν_b, $\nu = 1, 2,$

$b = 1, \ldots, n$, such that

$$\left| \left(\frac{\partial^2 S^\mu}{\partial z^c \, \partial a_b^\nu} = \frac{\partial \psi_c^\mu}{\partial a_b^\nu} \right) \right| \neq 0 \tag{4,45}$$

and for which the eqs. (4,6) are fulfilled. Then each of the $2n$ parameters a_b^ν will generate a current:

$$G_\nu^{\mu;b} := \frac{\partial S^\mu}{\partial a_b^\nu}(x, z; a) \,. \tag{4,46}$$

Suppose now, we have $4n$ functions $g_\nu^{\mu;b}(x)$, which are arbitrary, up to the following properties: They should obey the equations

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu} g_\nu^{\mu;b}(x) = 0 \tag{4,47}$$

identically for all x, b and ν and the equations

$$G_\nu^{\mu;b}(x, z; a) = \frac{\partial S^\mu}{\partial a_b^\nu}(x, z; a) = g_\nu^{\mu;b}(x) \tag{4,48}$$

should be a solvable system of $4n$ equations for the n variables z^b! Thus $3n$ of the eqs. (4,48) cannot be independent of the other n ones. Consider the following example: If the vector $y \in \mathrm{R}^n$ should be a solution of the two inhomogeneous linear systems $Ay = a$, $By = b$, where A and B are nonsingular $(n \times n)$-matrices, we must have $BA^{-1}a = b$.

Functions $g_\nu^{\mu;b}(x)$ with the property (4,47) are not difficult to find: Let $h_\nu^b(x)$ be $2n$ arbitrary smooth functions. If we define

$$g_\nu^{\mu;b}(x) = \varepsilon^{\mu\lambda} \partial_\lambda h_\nu^b(x) \,, \tag{4,49}$$

then the eqs. (4,47) are fulfilled identically.

Thus, the problem is, to find appropriate solutions $S^\mu(x, z; a)$ and functions $g_\nu^{\mu;b}(x)$ such that the $4n$ eqs. (4,48) do have n solutions $z^a = f^a(x)$. If such solutions $z^a = f^a(x)$ can be found, they are extremals; for it follows from eqs. (4,33), (4,48) and (4,47) that

$$(\partial_\mu z^b(x) - \partial H/\partial \pi_b^\mu)\, \partial \psi_b^\mu / \partial a_c^\nu = 0 \,, \qquad \nu = 1, 2 \,; \qquad c = 1, \ldots, n \,.$$

Combined with the inequality (4,45) we conclude

$$\partial_\mu z^b(x) - \partial H/\partial \pi_b^\mu = 0$$

and it follows from the construction that the integrability condition (4,6) and the eqs. (4,18) are fulfilled. It is clear that we will only find functions $g_\nu^{\mu;b}(x)$ with the properties required if the solutions $S^\mu(x, z)$ do obey the integrability conditions (4,6), at least locally!

Example:
The following is a rather simple example of a complete integral which allows for the construction of a solution of the field equation [von Rieth, 1982]:
If $V(z)=\frac{1}{2}m^2z^2$, then the HJ eq. (4,7) implies for the ansatz $S^\mu(x,z)=\frac{1}{2}h^\mu(x)z^2$ the Riccati-type equation

$$\partial_\mu h^\mu + h_\mu h^\mu + m^2 = 0\,,$$

which has the solution

$$h^0 = -\frac{1}{\sqrt{2}}m\tan\Big(\frac{1}{\sqrt{2}}mx^0 + a^0\Big)\,,\qquad h^1 = -\frac{1}{\sqrt{2}}m\tanh\Big(\frac{1}{\sqrt{2}}mx^1 + a^1\Big)\,,$$

with the in general nonsingular determinant

$$\Big|\Big(\frac{\partial S^\mu}{\partial a^\nu\,\partial z}\Big)\Big| = m^2z^2\Big/\Big\{2\cos^2\Big(\frac{1}{\sqrt{2}}mx^0 + a^0\Big)\cosh^2\Big(\frac{1}{\sqrt{2}}mx^1 + a^1\Big)\Big\}\,.$$

The quantities $G^\mu_\nu = \partial S^\mu/\partial a^\nu$ here have the form

$$G^0_0(x,z) = -mz^2\Big/\Big\{2\sqrt{2}\cos^2\Big(\frac{1}{\sqrt{2}}mx^0 + a^0\Big)\Big\}\,,\qquad G^1_0 = 0\,,$$

$$G^0_1 = 0\,,\qquad G^1_1(x,z) = -mz^2\Big/\Big\{2\sqrt{2}\cosh^2\Big(\frac{1}{\sqrt{2}}mx^1 + a^1\Big)\Big\}\,.$$

Taking for $g^\mu_\nu(x)$ the functions

$$g^0_0 = -\frac{1}{2\sqrt{2}}A^2m^3\cosh^2\Big(\frac{1}{\sqrt{2}}mx^1 + a^1\Big)\,,\qquad A = \text{const.}\,,\qquad g^1_0 = 0\,,$$

$$g^0_1 = 0\,,\qquad g^1_1 = -\frac{1}{2\sqrt{2}}A^2m^3\cos^2\Big(\frac{1}{\sqrt{2}}mx^0 + a^0\Big)\,,$$

we have $\partial_\mu g^\mu_\nu = 0$, as required. The equations $G^\mu_\nu(x,z) = g^\mu_\nu(x)$ have the solution

$$z(x) = \pm Am\cos\Big(\frac{1}{\sqrt{2}}mx^0 + a^0\Big)\cosh\Big(\frac{1}{\sqrt{2}}mx^1 + a^1\Big)\,,$$

which satisfies the KG equation.

4.4. A construction of wave fronts for a family of extremals depending on n parameters

Up to now we have been mainly concerned with the problem, whether a given solution $S^\mu(x,z)$ of the DWHJ eq. (4,3b) can be "generated" by extremals, i.e. whether the integrability conditions (4,6) are fulfilled. In most cases they are not. On the other hand, we saw in mechanics that a family of solutions $q^j(t) = f^j(t; u)$, $p_j(t) = g_j(t; u)$, depending on n parameters u^k such that the Lagrange brackets $[u^j, u^k]$ vanish and for which $|(\partial q^j/\partial u^k)| \neq 0$, generate the associated HJ function $S(t, q)$. Similarly one can ask, which wave fronts – if any – can be generated by a given family of extremals $z^a = f^a(x; u)$ which depend on n parameters u^b, $b = 1, \ldots, n$, such that

$$|(\partial z^a/\partial u^b)| \neq 0 . \tag{4,50}$$

This inequality allows one to solve the equations $z^a = f^a(x; u)$ for u:

$$u^a = \chi^a(x, z) . \tag{4,51}$$

If we insert the functions (4,51) into $\partial_\mu f^a(x; u)$:

$$\partial_\mu z^a(x) = \partial_\mu f^a[x; u = \chi(x, z)] = \varphi^a_\mu(x, z) ,$$

then the slope functions $\varphi^a_\mu(x, z)$ fulfill the integrability condition (4,6) by construction. The question arises whether these slope functions φ^a_μ can be generated by wave fronts, namely, whether there is a closed 2-form Ω of the differentials dx, dz^a, the coefficients of which are completely determined by the solutions $z^a = f^a(x)$. It is indeed possible to construct such a form Ω [Debever, 1937; see also Hilbert, 1906], however, it will in general not be of the DeDonder–Weyl type. We start with the general expression

$$\Omega = L\, dx^1 \wedge dx^2 + \pi^\mu_a\, dz^a \wedge d\Sigma_\mu + \tfrac{1}{2} h_{ab}\, \omega^a \wedge \omega^b .$$

Since $dz^a = \partial_\mu f^a(x)\, dx^\mu + (\partial f^a/\partial u^b)\, du^b$, we have $\omega^a = (\partial f^a/\partial u^b)\, du^b$, and therefore

$$\Omega = \Lambda(x, u)\, dx^1 \wedge dx^2 + l^\mu_a\, du^a \wedge d\Sigma_\mu + \tfrac{1}{4} K_{ab}\, du^a \wedge du^b , \tag{4,52}$$

where

$$\Lambda(x, u) = L[x, f(x; u), \partial_\mu f(x; u)] , \tag{4,53a}$$

$$l^\mu_a(x, u) = g^\mu_b \frac{\partial f^b}{\partial u^a} , \qquad g^\mu_b(x; u) = \pi^\mu_b , \tag{4,53b}$$

$$K_{ab}(x, u) = h_{rs}\left(\frac{\partial f^r}{\partial u^a}\frac{\partial f^s}{\partial u^b} - \frac{\partial f^r}{\partial u^b}\frac{\partial f^s}{\partial u^a}\right) . \tag{4,53c}$$

The postulate $d\Omega = 0$ implies the equations

$$\frac{\partial \Lambda}{\partial u^a} = \partial_\mu l^\mu_a , \tag{4,54}$$

$$\frac{\partial}{\partial u^a} l^\mu_b - \frac{\partial}{\partial u^b} l^\mu_a + \frac{1}{2} \varepsilon^{\mu\nu} \partial_\nu K_{ab} = 0 , \tag{4,55}$$

$$\frac{\partial}{\partial u^a} K_{bc} + \frac{\partial}{\partial u^c} K_{ab} + \frac{\partial}{\partial u^b} K_{ca} = 0 . \tag{4,56}$$

Because

$$\frac{\partial \Lambda}{\partial u^a} = \partial_b L \frac{\partial f^b}{\partial u^a} + \frac{\partial L}{\partial v^b_\mu} \frac{\partial}{\partial u^a}(\partial_\mu f^b) ,$$

$$\partial_\mu l^\mu_a = \partial_\mu g^\mu_b \frac{\partial f^b}{\partial u^a} + g^\mu_b\, \partial_\mu \left(\frac{\partial f^b}{\partial u^a}\right)$$

the eqs. (4,54) are a consequence of the Euler–Lagrange equations

$$\partial_\mu g^\mu_b - \partial_b L = 0 .$$

Furthermore, with

$$\frac{\partial}{\partial u^b} l^\mu_a - \frac{\partial}{\partial u^a} l^\mu_b = \frac{\partial f^c}{\partial u^a} \frac{\partial g^\mu_c}{\partial u^b} - \frac{\partial f^c}{\partial u^b} \frac{\partial g^\mu_c}{\partial u^a} =: [u^a, u^b]^\mu , \tag{4,57}$$

we can rewrite eqs. (4,55) as

$$\tfrac{1}{2}\varepsilon^{\mu\nu}\, \partial_\nu K_{ab} = [u^a, u^b]^\mu , \qquad \mu = 1, 2 . \tag{4,58}$$

The quantities $[u^a, u^b]^\mu$ represent generalizations of the Lagrange brackets (2,17) to field theories. The "field" eqs. (4,54) imply the "continuity" equation

$$\partial_\mu [u^a, u^b]^\mu = 0 , \tag{4,59}$$

which is an obvious generalization of eq. (2,18).

If the Lagrange brackets $[u^a, u^b]^\mu$ vanish, it follows from eqs. (4,58) that we may take $K_{ab} = 0$, or $h_{ab} = 0$. In this case the extremals $z^a = f^a(x; u)$ can be embedded in a DeDonder–Weyl geodesic field $S^\mu(x, z)$. However, the Lagrange brackets (4,57) will not vanish in general. In that case the eqs. (4,58) and (4,59) imply that the quantities K_{ab} can be calculated by the path-independent line integral

$$K_{ab}(x, u) = 2 \int [u^a, u^b]^\mu \, d\Sigma_\mu . \tag{4,60}$$

The functions (4,60) fulfill the eqs. (4,56) identically.

Having calculated the functions $K_{ab}(x, u)$, we can determine the coefficients $h_{ab} = \hat{h}_{ab}(x, u)$ from the linear eqs. (4,53c). Using the functions (4,51) we can calculate $\hat{h}_{ab}(x, \chi(x, z)) =: \tilde{h}_{ab}(x, z)$ and therefore

$$\begin{aligned} &\tilde{\Omega} = -\tilde{H}\, dx^1 \wedge dx^2 + \psi^\mu_a\, dz^a \wedge d\Sigma_\mu + \tfrac{1}{2}\tilde{h}_{ab}\, dz^a \wedge dz^b , \\ &\tilde{H}(x, z) = \psi^\mu_a \varphi^a_\mu - \tfrac{1}{2}\tilde{h}^{\mu\nu}_{ab} \varphi^a_\mu \varphi^b_\nu - \tilde{L}(x, z) , \\ &\qquad \tilde{h}^{\mu\nu}_{ab} = \varepsilon^{\mu\nu}\tilde{h}_{ab} , \qquad \tilde{L}(x, z) = \Lambda(x, u = \chi(x, z)) , \qquad \psi^\mu_a(x, z) = g^\mu_a(x; u = \chi(x, z)) . \end{aligned} \tag{4,61}$$

Next one has to determine the rank r of the form (4,61) and then find r functions $S^\rho(x, z)$ such that $\Omega = \mathrm{d}S^1 \wedge \mathrm{d}S^2 + \cdots$, where some of the functions $S^\rho(x, z)$ may be the coordinates x^μ themselves.

The procedure just described, which allows to calculate wave fronts associated with a given n-parametric family of extremals can be generalized as follows [Lepage, 1941]:

Suppose we are given slope functions $\varphi^a_\mu(x, z)$ which obey the integrability conditions (4,6) and for which the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu}\frac{\partial L}{\partial v^a_\mu}(x, z, v) = \partial_a L(x, z, v)$$

are fulfilled, if we replace the v^a_μ by the functions $\varphi^a_\mu(x, z)$. Inserting the forms $\omega^a = \mathrm{d}z^a - \varphi^a_\mu\, \mathrm{d}x^\mu$ and $v^a_\mu = \varphi^a_\mu(x, z)$ into

$$\Omega = L\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 + \pi^\mu_a\, \omega^a \wedge \mathrm{d}\Sigma_\mu + \tfrac{1}{2} h_{ab} \omega^a \wedge \omega^b$$

gives a form $\tilde{\Omega}$ the coefficients of which depend only on the variables x and z^a. The postulate $\mathrm{d}\tilde{\Omega} = 0$ implies partial differential equations for the functions $h_{ab} = \tilde{h}_{ab}(x, z)$ which can be solved, in principle. Thus, any arbitrary completely integrable set of slope functions φ^a_μ, which define the system

$$\partial_\mu z^a(x) = \varphi^a_\mu(x, z)$$

of first-order differential equations, the solutions $z^a = f^a(x)$ of which are extremals, i.e. for which the functions $f^a(x)$ are solutions of the Euler–Lagrange equations, too, can be generated by wave fronts.

4.5. The reduction of the field equations to ordinary canonical equations within the DWHJ framework

In our discussion above we have learnt about two essentially new possibilities for finding solutions of the field equations by means of solving the DWHJ eq. (4,3b): We can look for solutions $S^\mu(x, z)$ which fulfill the integrability conditions (4,6) and then we can either solve the first-order differential eqs. (4,5), or, if the solution $S^\mu(x, z)$ is a complete integral, we can try to solve the algebraic eqs. (4,48).

We shall now discuss a third method for solving the field equations which uses the HJ framework in order to transform the canonical partial differential equations for fields into ordinary differential equations which have the same form as the canonical equations in mechanics. The idea goes back to E. Hölder [1939] who used it first in the context of Carathéodory's HJ theory for fields. It was rephrased and generalized in terms of differential forms by Lepage, who applied it to the DeDonder–Weyl theory, too [1942, §21, 22]. Later Lepage's ideas were discussed – in more conventional language – by Van Hove [1945a] and by Klötzler [1970, §24, 25]. A modernized version was given by Goldschmidt and Sternberg [1973]. The following version is essentially that of Lepage.

The crucial starting point is again that the DWHJ eq. (4,3b) is one partial differential equation for two functions $S^\mu(x, z)$, so that we can choose one of them, e.g. $S^2(x, z)$, rather freely, provided the relation

$$\pi^2_a(x) = \partial_a S^2(x, z = f(x)) \tag{4,62}$$

is fulfilled on the extremals $z^a = f^a(x)$ under consideration. Let us suppose that we have somehow managed to find such a function S^2 – how this may be done will be briefly discussed below –. Consider next the canonical DeDonder–Weyl 2-form

$$\mathrm{d}S^1 \wedge \mathrm{d}x^2 + \mathrm{d}x^1 \wedge \mathrm{d}S^2 = -H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \pi_a^1\,\mathrm{d}z^a \wedge \mathrm{d}x^2 + \pi_a^2\,\mathrm{d}x^1 \wedge \mathrm{d}z^a\,. \tag{4,63}$$

Since $\mathrm{d}S^2 = \partial_\mu S^2\,\mathrm{d}x^\mu + \partial_a S^2\,\mathrm{d}z^a$, $\partial_a S^2 = \pi_a^2$, the form (4,63) can be reduced to

$$\mathrm{d}S^1 \wedge \mathrm{d}x^2 = (-\hat{H}\,\mathrm{d}x^1 + \pi_a^1\,\mathrm{d}z^a) \wedge \mathrm{d}x^2 =: \hat{\Omega}_0 \tag{4,64}$$

$$\hat{H} = H[x, z, \pi_a^1, \pi_a^2 = \partial_a S^2(x,z)] + \partial_{(2)} S^2(x,z)\,, \qquad \partial_{(2)} := \partial/\partial x^2\,.$$

We see that the 2-form (4,64) defines a 1-dimensional canonical system modulo $\mathrm{d}x^2$ in terms of the effective canonical 1-form $\hat{\theta} = -\hat{H}\,\mathrm{d}x^1 + \pi_a^1\,\mathrm{d}z^a$. The 2-form $\mathrm{d}\hat{\theta} = -\mathrm{d}H \wedge \mathrm{d}x^1 + \mathrm{d}\pi_a^1 \wedge \mathrm{d}z^a$ will contain terms involving $\mathrm{d}x^2$, since H and π_a^1 depend on x^2. However all these terms drop out of $\mathrm{d}\hat{\Omega}_0$ because of the overall factor $\mathrm{d}x^2$ in $\hat{\Omega}_0$ itself. Thus we have transformed the original 2-dimensional canonical system into a 1-dimensional one in the hyperplane $x^2 = \text{const.}$, with the effective canonical form

$$\hat{\theta} = -\hat{H}(t, z, p)\,\mathrm{d}t + p_a\,\mathrm{d}z^a\,, \qquad t = x^1\,, \qquad p_a = \pi_a^1\,, \tag{4,65}$$

which implies the canonical equations

$$\mathrm{d}p_a/\mathrm{d}t = -\partial_a \hat{H}\,, \qquad \mathrm{d}z^a/\mathrm{d}t = \partial\hat{H}/\partial p_a\,. \tag{4,66}$$

That the planes $x^2 = \text{const.}$ occur in this context is a consequence of the fact that these planes are "characteristic" planes of the form Ω_0 – see the discussion at the beginning of this chapter –.

The validity of the canonical eqs. (4,66) can be proven by direct calculation, too: Differentiating $\hat{H}$ with respect to z^a gives

$$\partial_a \hat{H} = \partial_a H + \frac{\partial H}{\partial \pi_b^2}\partial_b\partial_a S^2 + \partial_{(2)}\partial_a S^2(x,z) = \partial_a H + \frac{\mathrm{d}}{\mathrm{d}x^2}\partial_a S^2 = \partial_a H + \frac{\mathrm{d}\pi_a^2}{\mathrm{d}x^2} = -\frac{\mathrm{d}\pi_a^1}{\mathrm{d}x^1}\,,$$

because $\mathrm{d}\pi_a^\mu/\mathrm{d}x^\mu + \partial_a H = 0$. Furthermore, we have $\partial\hat{H}/\partial\pi_a^1 = \partial H/\partial\pi_a^1$. The above considerations may be rephrased in the following way: Assume that we have found a function $S^2(x,z)$ with the property (4,62). Then the DWHJ eq. (4,3b) becomes a partial differential equation of first order for the function $S^1(x,z)$:

$$\partial_{(1)} S^1 + \hat{H}(x, z, \pi_a^1 = \partial_a S^1) = 0\,. \tag{4,67}$$

The characteristic ordinary equations (with independent variable t) associated with this partial differential equation [Carathéodory, 1935, §43] are

$$\dot{x}^1 = 1\,, \qquad \dot{x}^2 = 0\,; \qquad \dot{x}^\mu := \mathrm{d}x^\mu/\mathrm{d}t\,, \tag{4,68a}$$

$$\dot{z}^a = \frac{\partial\hat{H}}{\partial\pi_a^1} = \frac{\partial H}{\partial\pi_a^1}\,, \qquad \dot{\pi}_a^1 = -\partial_a\hat{H} = -\partial_a H - \frac{\partial H}{\partial\pi_b^2}\partial_a\partial_b S^2 - \partial_a\partial_{(2)} S^2\,. \tag{4,68b}$$

From $\dot{x}^1 = 1$ we get $x^1 = t$ and the equation $\dot{x}^2 = 0$, which is a consequence of the fact that H does not contain $\partial_{(2)}S^1$, means that the variable x^2 has to be kept constant. The eqs. (4,68b) are the same as the eqs. (4,66).

The question arises, how we can recover the x^2-dependence of the fields $z^a = f^a(x)$, $\pi^\mu_a = g^\mu_a(x)$, after we have found solutions $z^a(t) = \hat{f}^a(t; u)$, $p_a(t) = \pi^1_a(t) = \hat{g}_a(t; u)$, where the parameters u^j, $j = 1, 2, \ldots$, are constants of integration. The x^2-dependence of the fields $f^a(x)$, $g^\mu_a(x)$ can be present in the functions $\hat{f}^a(t; u)$ and $\hat{g}_a(t; u)$ in 2 ways, "openly" and "hidden": In general the choice of the function $S^2(x, z)$ will be such that $\pi^2_a = \partial_a S^2(x, z)$ and $\partial_{(2)}S^2(x, z)$ will depend on x^2, and therefore x^2 will appear explicitly in H as a parameter which is to be treated as a constant when we solve the effective 1-dimensional problem with respect to the independent variable x^1. In order to keep the following formulae simple, we shall consider x^2 to be one of the parameters u^j. In addition the x^2-dependence may be hidden in the constants of integration $u^j \neq x^2$! This x^2-dependence can be determined by the following "variation of the constants" u^j: Knowing $\pi^1_a = \hat{g}_a(t; u)$, $\pi^2_a = \partial_a S^2(x, \hat{f}(t, u))$ we can calculate the slope functions

$$\varphi^a_2(x, z = \hat{f}(t, u)) = \hat{\varphi}^a_2(x, u)$$

and $\hat{f}^a(t; u)$ has to obey the differential equation

$$\partial_{(2)}\hat{f}^a(t, u) = \frac{\partial \hat{f}^a}{\partial u^j}\, \partial_{(2)}u^j = \hat{\varphi}^a_2(x, u)\,. \tag{4,69}$$

In addition we must have

$$\begin{aligned} \frac{\mathrm{d}\pi^2_a}{\mathrm{d}x^2} &= \frac{\mathrm{d}}{\mathrm{d}x^2}[\partial_a S^2(x, \hat{f}(t, u))] \\ &= \partial_{(2)}\partial_a S^2(x, \hat{f}) + (\partial_a\partial_b S^2)\frac{\partial \hat{f}^b}{\mathrm{d}u^j}\, \partial_{(2)}u^j \\ &= -\partial_a H[x, z = \hat{f}, \pi^1_a = \hat{g}_a, \pi^2_a = \partial_a S^2(x, \hat{f})] - \partial_{(1)}\hat{g}_a(t, u)\,. \end{aligned} \tag{4,70}$$

Let me illustrate these considerations by a simple example:
From eq. (4,15) we get

$$\partial_{(1)}S^1 = -\tfrac{1}{4}A, \qquad \pi^1 = \partial_z S^1 = kW'(z) = \frac{k}{\mu}(A - 2V(z))^{1/2}$$

and therefore for $\hat{H}$:

$$\hat{H} = H + \partial_{(1)}S^1 = \frac{1}{2}A\left(\frac{1}{2} - \frac{\omega^2}{\mu^2}\right) + \frac{1}{2}p^2 + \frac{\omega^2}{\mu^2}V(z)\,, \qquad p = \pi^0\,.$$

In this case π^1 and $\partial_{(1)}S^1$ and therefore $\hat{H}$, too, do not depend explicitly on x^1 or x^0. To be even more specific let me take $V(z) = \frac{1}{2}\mu^2 z^2$. Then we get – in the plane $x^2 = \text{const.}$ – the equations of motion

$$\mathrm{d}z/\mathrm{d}t = \partial\hat{H}/\partial p = p\,, \qquad \mathrm{d}p/\mathrm{d}t = -\partial\hat{H}/\partial z = -\omega^2 z\,,$$

with the solution $z = f(t; u) = u^1 \sin(\omega t + u^2)$. Since from the beginning – see eq. (4,14) –

$$H = \tfrac{1}{2}(\pi^0)^2 - \tfrac{1}{2}(\pi^1)^2 + V(z) = \tfrac{1}{2}A$$

we have $(u^1)^2 = A/\mu^2$. Thus there is no "hidden" x^2-dependence of u^1. For the function $\hat{\varphi}_1(x, u)$ we obtain

$$\hat{\varphi}_1 = -\frac{k}{\mu}(A - \mu^2 z^2)^{1/2} = -\frac{k}{\mu}\sqrt{A}\cos(\omega t + u^2)\,.$$

The eqs. (4,69) here take the form

$$\partial_1 \hat{f}(t; \mu) = \frac{\sqrt{A}}{\mu}\cos(\omega t + u^2)\, \partial_1 u^2 = -\frac{k}{\mu}\sqrt{A}\cos(\omega t + u^2)\,,$$

and have the solution $u^2 = -kx^1 + \text{const.}$

4.6. Embedding a given extremal into a system of wave fronts

The procedure just discussed can be used in order to embed any given extremal $z^a = f^a(x)$ weakly and locally in a geodesic field $S^\mu(x, z)$ of the DeDonder–Weyl type, such that the relations

$$\pi^\mu_a(x) = \partial_a S^\mu(x, z = f(x))$$

hold in the points $(x, f(x))$ of the extremal Σ^2_0, but it is not required that the integrability conditions (4,6) hold in a neighborhood of Σ^2_0.

We have seen above that we can reduce the problem to a 1-dimensional one, provided we can find functions $S^2(x, z)$ which fulfill the conditions (4,62) on Σ^2_0. For a given extremal $z^a = f^a(x)$ the condition (4,62) can be realized in different ways:

(i) Suppose the extremal $z^a = f^a_0(x)$ is a member of an n-parametric family of extremals $z^a = f^a(x, u)$ with $f^a_0(x) = f^a(x; u = u_0)$, such that $|(\partial z^a/\partial u^b)| \neq 0$, then we can solve the equations $z^a = f^a(x; u)$ for $u^a = \chi^a(x, z)$ and insert these functions into $\partial_\mu z^a(x, u = \chi(x, z)) = \varphi^a_\mu(x, z)$. If, in addition, the Lagrange brackets $[u^a, u^b]^\mu$ vanish, we can determine the functions $S^\mu(x, z)$ as discussed in section 4.4 above. If the Lagrange brackets do not vanish, we can calculate $\pi^\mu_a = \psi^\mu_a(x, z)$ from the slope functions φ^a_μ and solve the partial differential equations $\partial_a S^2(x, z) = \psi^2_a(x, z)$ for $S^2(x, z)$, in order to calculate the term $\partial_{(2)} S^2$ in the effective Hamiltonian $\hat{H}$. The procedure just described might be useful if the system is invariant under a transformation group depending on at least n parameters. In that case we can generate an n-parametric family of solutions by applying a general group element to $f^a_0(x)$. If the group is compact or contains compact subgroups the functions $u^a = \chi^a(x, z)$ will only exist locally!

(ii) Another method [Van Hove, 1945a], which is always applicable, uses a linear ansatz for $S^2(x, z)$ in a neighborhood of a point $(x, f(x))$:

$$S^2(x, z) = (z^a - f^a(x))\, \pi^2_a(x)\,. \tag{4,71}$$

This ansatz leads to

$$\partial_{(2)}S^2(x, z) = \partial_{(2)}\pi_a^2(x)\,(z^a - f^a(x)) - \partial_{(2)}f^a(x)\cdot \pi_a^2(x)\,.$$

Because

$$\hat{H} = H(x, z, p; \pi^2) + \partial_{(2)}S^2(x, z)\,,$$

the eqs. (4,66) become

$$\dot{z}^a = \partial H/\partial \pi_a^1\,, \qquad \dot{p}_a = -\partial_a H - \partial_{(2)}\pi_a^2(x)\,. \tag{4,72}$$

We see that the additional term $\partial_{(2)}\pi_a^2(x)$ represents an explicitly time-dependent "force" in the reduced problem.

Instead of pursuing the general case, let me discuss a simple example in order to illustrate the relevant points:

For $L = \frac{1}{2}(v_0) - \frac{1}{2}(v_1)^2 - V(z)$, we have $p^0 = v_0$, $p^1 = -v_1$, $H = \frac{1}{2}(p^0)^2 - \frac{1}{2}(p^1)^2 + V(z)$ and the DWHJ equation is

$$\partial_\mu S^\mu + \tfrac{1}{2}(\partial_z S^0)^2 - \tfrac{1}{2}(\partial_z S^1)^2 + V(z) = 0\,.$$

Let $z = f(x)$ be an extremal, i.e. a solution of the equation

$$\partial_0^2 f - \partial_1^2 f + V'(f) = 0\,. \tag{4,73}$$

The ansatz

$$S^1(x, z) = -\partial_1 f(x)\,(z - f(x))$$

gives

$$\partial_1 S^1(x, z) = -\partial_1^2 f(x)\,(z - f(x)) + (\partial_1 f(x))^2\,,$$

and therefore

$$\hat{H} = \tfrac{1}{2}p^2 + \tfrac{1}{2}[\partial_1 f(x)]^2 + V(z) - \partial_1^2 f(x)\,(z - f(x))\,. \tag{4,74}$$

The characteristic eqs. (4,72) take the form

$$\dot{z} = p\,, \qquad \dot{p} = -V'(z) + \partial_1^2 f(x)\,, \tag{4,75}$$

or

$$\ddot{z} = -V'(z) + \partial_1^2 f(t, x^2 = \text{const.})\,. \tag{4,76}$$

In general it will be difficult to solve the inhomogeneous eq. (4,76) explicitly. In the special case of $V = \frac{1}{2}\mu^2 z^2$ (Klein–Gordon equation) eq. (4,76) is that of a driven harmonic oscillator and can be solved analytically. But, already for the Sine–Gordon equation with $V(z)=\alpha(1-\cos\beta z)$ eq. (4,76) becomes that of a driven pendulum and in general will be hard to solve analytically.

However, as long as we are interested in wave fronts which "embed" a given extremal only locally, i.e. only in a small neighborhood of the surface $z=f(x)$, we can use the following quadratic approximation of $H(x,z,p)$ in the variable $\varepsilon(x,z):=z-f(x)$:

$$\hat{H}^{(2)}=\tfrac{1}{2}p^2+\tfrac{1}{2}(\partial_1 f)^2+V(f)+V'(f)\,\varepsilon+\tfrac{1}{2}V''(f)\,\varepsilon^2-\partial_1^2 f\varepsilon$$

which, because of the field eq. (4,73), reduces to

$$\hat{H}^{(2)}=\tfrac{1}{2}p^2+\tfrac{1}{2}(\partial_1 f)^2+V(f)+\tfrac{1}{2}V''(f)\,\varepsilon^2-\partial_0^2 f\varepsilon\,. \tag{4,77}$$

The characteristic eqs. (4,66) take the form

$$\dot{z}=\partial_0 f+\dot{\varepsilon}=\partial\hat{H}^{(2)}/\partial p=p\,,\qquad \dot{p}=-\partial\hat{H}^{(2)}/\partial z=\partial_0^2 f-V''(f)\,\varepsilon\,, \tag{4,78}$$

or

$$\ddot{\varepsilon}+V''(f)\,\varepsilon=0\,. \tag{4,79}$$

We now can construct a solution of the HJ equation

$$\partial_0 S^0+\hat{H}^{(2)}=0\,,\qquad p=\partial_z S^0\,, \tag{4,80}$$

without even solving eqs. (4,79) explicitly. The crucial point is that this differential equation is a homogeneous one: Let $g(t)$ be a solution with the initial condition $g(0)=1$, $\dot{g}(0)=0$. If ε_0 is a parameter $\neq 0$, then $\varepsilon(t)=\varepsilon_0\,g(t)$ is a solution of eq. (4,79) with $\varepsilon(0)=\varepsilon_0$, $\dot{\varepsilon}(0)=0$ and we have $\dot{\varepsilon}(t)=\varepsilon_0\dot{g}=\varepsilon(t)\dot{g}/g$. Direct integration of the differential equation

$$\begin{aligned}\dot{\sigma}&=p(t)\frac{\partial\hat{H}^{(2)}}{\partial p}[x,z(t),p(t)]-\hat{H}^{(2)}[x,z(t),p(t)]\\&=\tfrac{1}{2}(\partial_t f+\dot{\varepsilon})^2-\tfrac{1}{2}(\partial_1 f)^2-V(f)+\partial_t^2 f\,\varepsilon(t)-\tfrac{1}{2}V''(f)\,\varepsilon^2(t)\end{aligned}$$

yields the solution

$$\sigma(x^0;x^1,\varepsilon)=\text{const.}+\partial_0 f\varepsilon-(\partial_t f\varepsilon)_{x^0=0}+\tfrac{1}{2}\varepsilon\dot{\varepsilon}+\int_0^{x_0}\mathrm{d}t\,[\tfrac{1}{2}(\partial_t f)^2-\tfrac{1}{2}(\partial_1 f)^2-V(f(t,x^1))]\,, \tag{4,81}$$

where partial integration, eq. (4,79) and $\dot{\varepsilon}(0)=0$ has been used. Choosing the constant in eq. (4,81) to be equal to $(\partial_t f\varepsilon)_{x^0=0}$ and recalling that $\dot{\varepsilon}=\varepsilon\dot{g}/g$ we obtain the solution

$$\begin{aligned}S^0(x,z)&=\sigma(x^0;x^1,\varepsilon=z-f(x))\\&=\partial_0 f(x)\,(z-f(x))+\tfrac{1}{2}(\partial_0 g/g)(z-f(x))^2+\int_0^{x^0}\mathrm{d}t\,L[f(t,x^1),\partial_\mu f(t,x^1)]\end{aligned} \tag{4,82}$$

of eq. (4,80). Notice that $\partial_z S^0|_{\varepsilon=0}=\partial_0 f(x)$, as it should.

As to the integrability condition (4,9) we have

$$-\partial_z^2 S^0\,\partial_z S^1 + \partial_1\partial_z S^0 = \partial_1\partial_0 f + (z - f(x))\,\partial_1(\partial_0 g/g)\,,$$
$$-\partial_z^2 S^1\,\partial_z S^0 - \partial_0\partial_z S^1 = \partial_1\partial_0 f\,,$$

i.e. the conditions (4,9) are fulfilled if either $z = f(x)$, i.e. on the extremal, or if $\partial_1(\partial_0 g/g) = 0$, i.e. if $V''(f(x^0, x^1))$ does not depend on x^1, which is the case for $V(z) = \frac{1}{2}\mu^2 z^2$.

Suppose now the extremal $z = f(x)$ depends on a parameter a. Then the components G^0, G^1 of the associated DWHJ current are

$$G^0(x) = \left.\frac{\partial S^0}{\partial a}\right|_{\varepsilon=0} = -\partial_0 f\frac{\partial}{\partial a}f + \frac{\partial}{\partial a}\int_0^{x^0} \mathrm{d}t\, L[f(t, x^1; a), \partial_\mu f(t, x^1; a)]\,,$$
$$G^1(x) = \left.\frac{\partial S^1}{\partial a}\right|_{\varepsilon=0} = \partial_1 f\frac{\partial}{\partial a}f\,. \tag{4,83}$$

Using the field eq. (4,73) one verifies immediately that $\partial_0 G^0 + \partial_1 G^1 = 0$.

Remarks:

(i) The generalization of the current (4,83) to arbitrary dimensions m and any number of fields is [von Rieth, 1982]:

$$G^0 = -\pi^0_b\frac{\partial}{\partial a}f^b + \frac{\partial}{\partial a}\int_0^{x^0}\mathrm{d}t\, L[f(t, \boldsymbol{x}; a), \partial_\mu f(t, \boldsymbol{x}; a)]\,,$$
$$G^{\bar\mu} = -\pi^{\bar\mu}_b\frac{\partial}{\partial a}f^b\,, \qquad \bar\mu = 1, \dots, m-1\,; \qquad \boldsymbol{x} = (x^1, \dots, x^{m-1})\,, \tag{4,84}$$

for which the continuity equation $\partial_\mu G^\mu = 0$ is again a consequence of the Euler–Lagrange equations.

(ii) If, for instance, a is the parameter τ of time translations $x^0 \to x^0 + \tau$, $\boldsymbol{x} \to \boldsymbol{x}$, then, if L does not depend explicitly on x^0, this parameter appears in the form $f^b(x^0 + \tau, \boldsymbol{x})$ and the current (4,84) becomes

$$G^\mu = \delta^\mu_0 L - \pi^\mu_b\,\partial_0 f^b = -T^\mu_0\,,$$

which is the usual energy current! More generally, the expression (4,84) for a conserved current contains the Noether current as a special case, for which the parameter a becomes a group parameter of a symmetry transformation. There is, however, an important difference between the currents associated with a group parameter and those HJ currents not associated with a general symmetry transformation: In the case of symmetry transformations the quantities $\partial f^b/\partial a$, $\partial\partial_\mu f^b/\partial a$ are again expressible in terms of the quantities f^b and $\partial_\mu f^b$. This will in general not be possible for arbitrary parameters.

(iii) In the current (4,84) the zero component seems to play a special role. This is not so: We get a conserved current, if an arbitrary component G^ν has the form

$$G^\nu = \frac{\partial}{\partial a}\int_0^{x^\nu}\mathrm{d}\bar x^\nu\, L[f(x^0, \dots, \bar x^\nu, \dots, x^{m-1}), \partial_\mu f(x^0, \dots, \bar x^\nu, \dots, x^{m-1})] - \pi^\nu_b\frac{\partial}{\partial a}f^b\,,$$

and the rest is given by

$$G^\mu = -\pi^\mu_b \frac{\partial}{\partial a} f^b, \qquad \mu \neq \nu.$$

Another interesting solution, due to Rinke [1981], of the DWHJ equation in the case of the relativistic string [Nambu, 1970; Rebbi, 1974; Scherk, 1975] is the following: For this system the dependent variables z^a are the coordinates x^α, $\alpha = 0, 1, 2, 3$, of the Minkowski space M^4 (with metric $x \cdot x = g_{\alpha\beta} x^\alpha x^\beta = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$) which depend on the parameters t^μ, $\mu = 1, 2$. The Lagrangian is

$$L = [-\tfrac{1}{2} v_{\alpha\beta} v^{\alpha\beta}]^{1/2} = [(v_1 \cdot v_2)^2 - (v_1 \cdot v_1)(v_2 \cdot v_2)]^{1/2}, \qquad v^{\alpha\beta} = v_1^\alpha v_2^\beta - v_1^\beta v_2^\alpha, \tag{4,85}$$

from which one obtains the canonical momenta

$$\begin{aligned} \pi^1_\alpha &= \frac{1}{L} g_{\alpha\beta} [v_2^\beta (v_1 \cdot v_2) - v_1^\beta (v_2 \cdot v_2)], \\ \pi^2_\alpha &= \frac{1}{L} g_{\alpha\beta} [v_1^\beta (v_1 \cdot v_2) - v_2^\beta (v_1 \cdot v_1)], \end{aligned} \tag{4,86}$$

with the properties

$$v_\mu \cdot \pi^\nu = \delta^\nu_\mu L \tag{4,87}$$

and

$$\pi_{\alpha\beta} := \pi^1_\alpha \pi^2_\beta - \pi^1_\beta \pi^2_\alpha = -g_{\alpha\gamma} g_{\beta\delta} v^{\gamma\delta}. \tag{4,88}$$

For the Hamilton function $H = \pi^\mu_\alpha v^\alpha_\mu - L$ we obtain, with the help of eqs. (4,87) and (4,88)

$$H(\pi) = L = (-\tfrac{1}{2} \pi_{\alpha\beta} \pi^{\alpha\beta})^{1/2}. \tag{4,89}$$

If $x^\alpha = f^\alpha(t)$, $t = (t^1, t^2)$ is a solution of the Euler–Lagrange equations $\partial_{(\mu)} \pi^\mu_\alpha = 0$, $\partial_{(\mu)} := \partial / \partial t^\mu$ then

$$S^\mu(t, x) = \pi^\mu_\alpha(t) (x^\alpha - \tfrac{1}{2} f^\alpha(t)) \tag{4,90}$$

is a solution of the DWHJ equation

$$\partial_{(\mu)} S^\mu + H(\pi) = 0, \qquad \pi^\mu_\alpha = \partial_\alpha S^\mu, \qquad \partial_\alpha := \partial / \partial x^\alpha. \tag{4,91}$$

The assertion follows immediately from the relations (4,87) and the field equations $\partial_{(\mu)} \pi^\mu_\alpha = 0$ which imply $\partial_{(\mu)} S^\mu = -L = -H$!

If $f^\alpha(t; a)$ is a solution of the field equations depending on a parameter a then the solutions (4,90) give the following components of a conserved DWHJ current

$$G^\mu(t) = \frac{\partial S^\mu}{\partial a}\bigg|_{x^\alpha = f^\alpha(t)} = \frac{1}{2}\left(\frac{\partial}{\partial a}\pi^\mu_\alpha\right) f^\alpha - \frac{1}{2}\pi^\mu_\alpha \frac{\partial f^\alpha}{\partial a}. \tag{4,92}$$

I finally mention an interesting relation between solutions $S^\mu(x, z)$ of the HJ eq. (4,7) *and* the integrability condition (4,9) and the solutions $z = f(x)$ of the field equation $(\partial_0^2 - \partial_1^2)z(x) + V'(z) = 0$ [Nüsser, 1982]: Differentiating the HJ eq. (4,7) with respect to z and combining the result

$$\partial_\mu \psi^\mu + \tfrac{1}{2}\partial_z[(\psi^0)^2 - (\psi^1)^2] = -V'(z), \qquad \psi^\mu := \partial_z S^\mu,$$

with the integrability condition (4,9), by adding and subtracting, gives the equations

$$(\partial_0 + \partial_1)(\psi^0 + \psi^1) + (\psi^0 - \psi^1)\,\partial_z(\psi^0 + \psi^1) = -V'(z), \tag{4,93a}$$

$$(\partial_0 - \partial_1)(\psi^0 - \psi^1) + (\psi^0 + \psi^1)\,\partial_z(\psi^0 - \psi^1) = -V'(z). \tag{4,93b}$$

If we have solutions $\psi^\mu(x, z)$ of the eqs. (4,93), then we get solutions $S^\mu(x, z)$ of the HJ eq. (4,7) by integration:

$$S^\mu(x, z) = \int_{z_0}^{z} \mathrm{d}\bar{z}\, \psi^\mu(x, \bar{z}) + s^\mu(x), \tag{4,94}$$

where the functions $s^\mu(x)$ are independent of z and which, according to the HJ eq. (4,7), have to obey the equation

$$\partial_\mu s^\mu(x) = -\tfrac{1}{2}\psi^\mu \psi_\mu - V(z) - \int_{z_0}^{z} \mathrm{d}\bar{z}\, \partial_\mu \psi^\mu. \tag{4,95}$$

As the l.h. side of the last equation is independent of z, we have

$$\partial_\mu \psi^\mu = -V'(z) - \tfrac{1}{2}\partial_z(\psi^\mu \psi_\mu),$$

and therefore eq. (4,95) takes the form

$$\partial_\mu s^\mu(x) = -V(z_0) - \tfrac{1}{2}\psi^\mu \psi_\mu(x, z_0), \tag{4,96}$$

from which $s^\mu(x)$ is to be determined, once the functions ψ^μ are known.

Introducing light cone variables $x_+ = \frac{1}{2}(x^0 + x^1)$, $x_- = \frac{1}{2}(x^0 - x^1)$ and the functions $a(x, z) = \psi^0 + \psi^1$, $b(x, z) = \psi^0 - \psi^1$, the eqs. (4,93) may be rewritten as

$$\partial_+ a + b\,\partial_z a = -V'(z) \tag{4,97a}$$

$$\partial_- b + a\,\partial_z b = -V'(z). \tag{4,97b}$$

The eqs. (4,97) have a very special structure: The first equation contains only derivatives with respect to x_+ and z, whereas the second one contains only derivatives with respect to x_- and z. Suppose we would know the function $b(x, z)$, then eq. (4,97a) yields the characteristic equations

$$\partial_s x_+ = 1\,, \qquad \partial_s x_- = 0\,, \qquad \partial_s z = b\,, \qquad \partial_s a = -V'(z) \tag{4,98}$$

for the 1-dimensional characteristic curves $x_+(s)$, $x_-(s)$ and $z(x)$, $s \in \mathrm{R}$. Correspondingly eq. (4,97b) implies the characteristic equations with respect to a second curve variable t:

$$\partial_t x_+ = 0\,, \qquad \partial_t x_- = 1\,, \qquad \partial_t z = a\,, \qquad \partial_t b = -V'(z)\,. \tag{4,99}$$

The first two of the eqs. (4,98) and (4,99) show that we may take $s = x_+$ and $t = x_-$. Then the rest of the equations implies the compatibility condition

$$\partial_+\partial_- z(x) + V'(z) = 0\,,$$

which is just the Euler–Lagrange equation!

If $z = f(x; u)$ is a solution of this field equation depending on a parameter u, such that $z = f(x; u)$ can be solved for u, i.e. $u = \chi(x, z)$, then

$$a(x, z) = \partial_- f[x; u = \chi(x, z)]\,, \qquad b(x, z) = \partial_+ f[x; u = \chi(x, z)]$$

are solutions of eqs. (4,97) and the functions $S^\mu(x, z)$ may be constructed as indicated above.

One may ask: what is the use of embedding a given extremal in a system of wave fronts? There are – at least – two physical reasons: First, we have seen that any family of wave fronts $S^\mu(x, z)$, which depends on a parameter, generates a conserved current "along" the associated extremals. Thus, embedding a given extremal may lead to interesting continuity equations associated with that special extremal, or a set of extremals. Second, HJ wave fronts, transversal to a family of extremals in mechanics, can provide useful semiclassical approximations for quantum mechanical problems [Berry and Mount, 1972; Voros, 1976; DeWitt-Morette, Maheshwari and Nelson, 1979]. HJ wave fronts associated with classical solutions of field equations may be equally useful for semiclassical approximations in quantum field theory.

4.7. Bibliographical notes

The early history of the HJ theory for fields has already been sketched in the Introduction. Independent of DeDonder's paper [1913] Prange's dissertation appeared in 1915, in which he attempted to combine ideas around Hilbert's independent integral with those of Volterra and Fréchet. In 1933 Born tried to apply Prange's investigation to the quantization of the electromagnetic field, without much success. Born's paper, however, stimulated Weyl to discuss the ansatz of DeDonder thoroughly [1934 and 1935]. In 1930 DeDonder himself had worked out his version of a HJ theory for fields. A second edition of this monograph appeared in 1935. For further references see the notes at the end of chapter 3, especially Lepage, Boerner, Dedecker, Sternberg and Goldschmidt. The most extensive recent treatment of the DWHJ theory can be found in the textbook by Klötzler [1970]. It is discussed in the textbooks by Funk [1970, ch. VI] and Rund [1973], too.

5. Carathéodory's canonical theory for fields

Probably the most interesting canonical theory for fields is that of Carathéodory. It has a rich geometrical structure and it is the only canonical theory for fields which has n-dimensional transversal wave fronts if the system has n dependent variables, in complete analogy to the situation in mechanics! However, one has to pay a prize for this structural richness: The canonical equations of motion and the CHJ equation are rather complicated which makes it difficult to handle them analytically. On the other hand, Carathéodory's canonical equations have the remarkable property that they can be cast into "mechanical" form by a canonical point transformation. This allows for an interesting reformulation of the problem to find solutions of the Euler–Lagrange equations! A very intriguing feature of Carathéodory's theory consists in its property to provide a number of new handles for the qualitative analysis of a system under consideration (singularities of families of solutions, bifurcations, singularities in transversality relations, caustics etc.). Some of these aspects will be discussed in chapter 8.

In any case, it is the main purpose of this review to draw Carathéodory's canonical theory to light, in the hope that a joint effort of the theoretical physics community will make it fruitful for physics at large!

5.1. The generalized Legendre transform

The basic defining relation for the canonical quantities in Carathéodory's theory is–see eq. (3,60)–:

$$\Omega_c = \frac{1}{L} a^1 \wedge a^2 = -\frac{1}{H_c} \theta^1 \wedge \theta^2 , \tag{5,1}$$

where

$$a^\mu = L\,dx^\mu + \pi^\mu_a \omega^a = -T^\mu_\nu\,dx^\nu + \pi^\mu_a\,dz^a , \qquad T^\mu_\nu = \pi^\mu_a v^a_\nu - \delta^\mu_\nu L , \tag{5,2}$$

and

$$\theta^\mu = -H_c\,dx^\mu + p^\mu_a\,dz^a = -S^\mu_\nu\,dx^\nu + p^\mu_a\,\omega^a , \qquad S^\mu_\nu = p^\mu_a\,v^a_\nu - \delta^\mu_\nu\,H_c . \tag{5,3}$$

On the extremals (T^μ_ν) is the canonical energy-momentum tensor!

Expanding the l.h. side of eq. (5,1) in terms of dx^μ and dz^a and comparing coefficients on both sides gives

$$H_c = -\frac{1}{L}|T| , \qquad |T| := \det(T^\mu_\nu) , \tag{5,4}$$

and

$$p^1_a = -\frac{1}{L}(\pi^1_a T^2_2 - \pi^2_a T^1_2) , \qquad p^2_a = -\frac{1}{L}(\pi^2_a T^1_1 - \pi^1_a T^2_1) .$$

Here it is convenient to use the notion of the algebraic complement (or cofactor) $\bar{T}^\nu_\mu$ of the matrix element T^μ_ν [see e.g. Satake, 1975, ch. II, §3], which in our special case form the matrix

$$\bar{T} = (\bar{T}^{\mu}_{\nu}) = \begin{pmatrix} T^2_2 & -T^1_2 \\ -T^2_1 & T^1_1 \end{pmatrix}$$

and which – like $n \times n$ matrices in general – obey the equations

$$T^{\mu}_{\rho}\bar{T}^{\rho}_{\nu} = \delta^{\mu}_{\nu}|T| = \bar{T}^{\mu}_{\rho}T^{\rho}_{\nu}\,. \tag{5,5}$$

They imply

$$|T|\,|\bar{T}| = |T|^m\,, \tag{5,6}$$

if (T^{μ}_{ν}) is an $m \times m$ matrix. With the help of the matrix $(\bar{T}^{\mu}_{\nu})$ we can rewrite the above expressions for the momenta p^{μ}_{a}:

$$p^{\mu}_{a} = -\frac{1}{L}\bar{T}^{\mu}_{\rho}\,\pi^{\rho}_{a}\,, \tag{5,7}$$

or, because of the relations (5,4) and (5,5)

$$H_c\pi^{\mu}_{a} = T^{\mu}_{\rho}p^{\rho}_{a}\,. \tag{5,8}$$

The last equation implies $H_c\, a^{\mu} = T^{\mu}_{\rho}\,\theta^{\rho}$.

In the following it will be convenient to use the matrices

$v = (v^{a}_{\mu})$: n rows, 2 columns,

$\pi = (\pi^{\mu}_{a}), p = (p^{\mu}_{a})$: 2 rows, n columns,

which allow us to write

$$T = \pi v - LE_2\,, \qquad p = -\frac{1}{L}\bar{T}\pi\,, \tag{5,9}$$

$$T\bar{T} = \bar{T}T = |T|\,E_2\,.$$

(Notice that our definition of $\bar{T}$ includes the interchange of the indices: $\overline{(T^{\mu}_{\nu})} = (\bar{T}^{\nu}_{\mu})$, which is convenient, if one employs the summation convention for coinciding upper and lower indices.) Finally, comparing the factors of $\mathrm{d}z^a \wedge \mathrm{d}z^b$ in eq. (5,1) we obtain

$$\begin{aligned} h_{ab} &= \frac{1}{L}(\pi^1_a\pi^2_b - \pi^1_b\pi^2_a) \\ &= -\frac{1}{H_c}(p^1_a p^1_b - p^1_b p^2_a) =: \eta_{ab}\,. \end{aligned} \tag{5,10}$$

If we use the basis vectors $\mathrm{d}x^{\mu}$ and ω^a for expressing the forms a^{μ} and θ^{μ} – see eqs. (5,2) and (5,3) –, we get from eq. (5,1) the relations

$$L = -\frac{1}{H_c}|S|\,, \qquad H_c\pi = -\bar{S}p\,, \tag{5,11}$$

and we have

$$S = pv - H_cE_2 = -\frac{1}{L}\bar{T}\pi v - H_cE_2$$

$$= -\frac{1}{L}\bar{T}(T + LE_2) - H_cE_2 = -\bar{T}\,, \tag{5,12}$$

where the relations (5,7), (5,9) and (5,4) have been used. We further notice that

$$Lp = S\pi\,. \tag{5,13}$$

The expression (5,4) for the Hamilton-function is intuitively extremely appealing: H_c is essentially the determinant of the canonical energy-momentum tensor $T = \pi v - LE_2$. It is very plausible physically that this quantity governs the dynamics of the system locally! Furthermore, since

$$|T| = L^2 - \mathrm{tr}(t^\mu_\nu)\,L + \det(t^\mu_\nu)\,, \qquad t^\mu_\nu := \pi^\mu_a v^a_\nu\,, \tag{5,14}$$

we see that $|T|$ can be looked at as the characteristic polynomial of the matrix $(t^\mu_\nu) = (\pi^\mu_a v^a_\nu)$, with the Lagrangian L as the usual polynomial variable λ, which here, however, is a function of v^a_μ etc., too. If we compare eq. (5,14) with the DeDonder–Weyl Hamilton-function

$$H_{\mathrm{DW}} = K = \pi^\mu_a v^a_\mu - L = \mathrm{tr}(t^\mu_\nu) - L\,, \tag{5,15}$$

we see that the function K contains only one invariant of the matrix (t^μ_ν), namely $\mathrm{tr}(t^\mu_\nu)$, whereas $H_c = -|T|/L$ contains both, $\mathrm{tr}(t^\mu_\nu)$ and $\det(t^\mu_\nu)$. (In the case of m independent variables H_c depends on the m invariants of the matrix $(t^\mu_\nu = \pi^\mu_a v^a_\nu)$). Thus, Carathéodory's canonical theory makes full use of all the invariants of the matrix $(\pi^\mu_a v^a_\nu)$!

It is very important that the DW theory can be looked at as the zero order approximation [Weyl, 1935], if one expands Carathéodory's canonical quantities in powers of $(1/L)$:

$$H_c = -L + \pi^\mu_a v^a_\mu - \frac{1}{L}|(t^\mu_\nu)| = H_{\mathrm{DW}} - \frac{1}{L}|(t^\mu_\nu)|\,,$$

$$p^1_a = \pi^1_a - \frac{1}{L}(\pi^1_a\pi^1_b v^b_2 - \pi^2_a\pi^1_b v^b_2)\,, \qquad \text{etc}\,.$$

In the language of physicists this means that *the standard canonical theory is the "strong coupling" limit*(!) *of Carathéodory's theory*: Suppose the Lagrangian L has the form $L = L_0(v) - V(z)$, where the interaction term $V(z)$ does not depend on v, so that $\pi^\mu_a = \partial L_0/\partial v^a_\mu$. Then the matrix $(\pi^\mu_a v^a_\nu)$ does not depend on $V(z)$ explicitly. So, if $|V| \gg |L_0|$, then $H_c \approx H_{DW}$ and $p^\mu_a \approx \pi^\mu_a$. This property remains valid for arbitrary dimensions m. Thus the conventional canonical theory appears to be an approximation of a more sophisticated one!

It is remarkable that the Hamilton function H_c can be expressed by the determinant $|R|$ of the matrix

$$R = (R^a_b = v^a_\mu \pi^\mu_b - \delta^a_b\, L) = v \cdot \pi - LE_n\,, \tag{5,16}$$

too.

Proof: It follows from the relations (2,40) that

$$\begin{aligned} |T| &= |-LE_2 + \pi \cdot v| = \left| \begin{pmatrix} E_n & v \\ -\pi & -LE_2 \end{pmatrix} \right| \\ &= (-L)^2 \left| E_n - \frac{1}{L} v \cdot \pi \right| = (-L)^{2-n} |R|\,, \end{aligned} \tag{5,17}$$

and therefore we have from eq. (5,4):

$$H_c = (-L)^{1-n} |R|\,. \tag{5,18}$$

Eq. (5,17) shows that the characteristic polynomial of the matrix $(\pi^\mu_a v^a_\nu)$ is proportional to that of the matrix $(v^a_\mu \pi^\mu_b)$! Notice that R is a matrix with respect to the field ("internal") indices a, whereas T is a matrix with respect to the "space-time" indices μ. From eqs. (5,8) and (5,13) we obtain

$$p \cdot R = p(v \cdot \pi - LE_n) = (p \cdot v)\, \pi - pL = (S + H_c E_2)\, \pi - pL = H_c \pi\,. \tag{5,19}$$

In analogy to the matrix S we define the matrix

$$Q = (Q^a_b = v^a_\mu p^\mu_b - \delta^a_b H_c) = v \cdot p - H_c E_n\,, \tag{5,20}$$

which has the property

$$R \cdot Q = LH_c E_n\,, \tag{5,21}$$

because

$$\begin{aligned} (v \cdot \pi - LE_n)(v \cdot p - H_c E_n) &= v(\pi \cdot v)\, p - Lv \cdot p - H_c v \cdot \pi + LH_c E_n \\ &= v(T + LE_2)\, p - Lv \cdot p - H_c v \cdot \pi + LH_c E_n = LH_c E_n\,. \end{aligned}$$

It follows from eqs. (5,21) and (5,18) that

$$|Q| = (-1)^{n-1}\, LH_c^{n-1}\,, \tag{5,22}$$

and from eqs. (5,21) and (5,20) that

$$Lp = \pi \cdot Q\,. \tag{5,23}$$

If $n = 1$, then $R = \pi^\mu v_\mu - L$, and we see from eq. (5,18) that $H_c = \pi^\mu v_\mu - L = K = H_{DW}$, and from eq. (5,23) that $p^\mu = \pi^\mu$, i.e. for $n = 1$ the canonical theories of DeDonder–Weyl and Carathéodory coincide! This is a consequence of the fact, mentioned already in chapter 3, that a 2-form in 3 variables – here x^1, x^2 and z – has rank 2.

In order to establish the relations between the quantities $\partial_\mu H_c$, $\partial_a H_c$, $\partial H_c/\partial p^\mu_a$ and $\partial_\mu L$, $\partial_a L$, v^a_μ, induced by the Legendre transformation $v \to p$, $L \to H_c$, we could use the relations (3,14). It is, however, more instructive, to derive them directly: In order to do so, we recall the relation [see, e.g., Satake, 1975, pp. 98–99]

$$\mathrm{d}|A| = \bar{A}^\nu_\mu \,\mathrm{d}A^\mu_\nu \tag{5,24}$$

for the differential of the determinant $|A|$ of a matrix $A = (A^\mu_\nu)$, where $\bar{A}^\nu_\mu$ is the cofactor of A^μ_ν. Applying the formula (5,24) to $H_c L = -|S|$, eq. (5,11), gives

$$H_c \,\mathrm{d}L + L \,\mathrm{d}H_c = -\bar{S}^\mu_\nu \,\mathrm{d}S^\nu_\mu = T^\mu_\nu \,\mathrm{d}S^\nu_\mu \,.$$

Because

$$\mathrm{d}S^\nu_\mu = v^a_\mu \,\mathrm{d}p^\nu_a + p^\nu_a \,\mathrm{d}v^a_\mu - \delta^\nu_\mu \,\mathrm{d}H_c\,, \qquad T^\mu_\nu p^\nu_a = H_c \pi^\mu_a\,,$$

we obtain the basic relation

$$K\Big(\mathrm{d}H_c - \frac{1}{K} v^a_\mu T^\mu_\nu \,\mathrm{d}p^\nu_a\Big) = -H_c(\mathrm{d}L - \pi^\mu_a \,\mathrm{d}v^a_\mu)\,, \qquad K = \pi^\mu_a v^a_\mu - L\,, \tag{5,25}$$

which implies

$$K\,\partial_\mu H_c = -H_c\,\partial_\mu L\,, \tag{5,26a}$$

$$K\,\partial_a H_c = -H_c\,\partial_a L\,, \tag{5,26b}$$

$$\frac{\partial H_c}{\partial p^\mu_a} =: w^a_\mu(x, z, p) = \frac{1}{K} v^a_\nu T^\nu_\mu = \frac{1}{K} R^a_b v^b_\mu\,. \tag{5,26c}$$

In matrix notation the last equation reads: $w = v \cdot T/K = R \cdot v/K$. Eq. (5,26c) expresses the quantity $v^a_\nu T^\nu_\mu/K$ as a function $w^a_\mu(x, z, p)$ of x, z and p. For $n = 1$ we have $v_\mu T^\mu_\nu/K = v_\nu = w_\nu(x, z, p)$. Combining eq. (5,26c) with eqs. (5,8) and (5,19) gives

$$\pi \cdot v/K = p \cdot w/H_c\,, \qquad w = (w^a_\mu)\,, \tag{5,27a}$$

$$v \cdot \pi/K = w \cdot p/H_c\,. \tag{5,27b}$$

Suppose we can solve the equations $w^a_\mu = \partial H_c/\partial p^\mu_a(x, z, p)$ for $p^\mu_a = p^\mu_a(x, z, w)$. We then can define the "normal" Legendre transform

$$G(x, z, w) = w^a_\mu p^\mu_a - H_c\,, \tag{5,28}$$

for which we get from the relations (5,27)

$$G/H_c = L/K. \tag{5,29}$$

For $n = 1$, we have $G = L$. The Legendre transformation $p \to w$, $H \to G$ defined by eq. (5,28) implies

$$\partial_\mu G = -\partial_\mu H_c, \qquad \partial_a G = -\partial_a H_c, \qquad \partial G/\partial w^a_\mu = p^\mu_a. \tag{5,30}$$

With the help of the eqs. (5,27) we can express the matrices T and R in terms of x, z, p and $w(x, z, p)$:

$$T/K = V/H_c, \qquad V = p \cdot w - GE_2, \tag{5,31a}$$

$$R/K = W/H_c, \qquad W = w \cdot p - GE_n. \tag{5,31b}$$

Eqs. (5,29) and (5,31) can be used in order to express L and K as functions of x, z and p: From eqs. (5,31) we obtain

$$|T|K^{-2} = |V|H_c^{-2}, \qquad |R|K^{-n} = |W|H_c^{-n}, \tag{5,32}$$

which, combined with the eqs. (5,4), (5,18) and (5,29), give

$$L = -H_c G^2/|V| = (-1)^{n-1} H_c G^n/|W|, \tag{5,33a}$$

$$K = -GH_c^2/|V| = (-1)^{n-1} H_c^2 G^{n-1}/|W|. \tag{5,33b}$$

It follows from eqs. (5,26c) and (5,31) that

$$v \cdot V = H_c w, \tag{5,34a}$$

$$W \cdot v = H_c w, \tag{5,34b}$$

from which we obtain

$$\begin{aligned} v^a_\mu &= K(w^a_\mu + g^{ab}_{\mu\nu} p^\nu_b)/H_c =: \hat{\varphi}^a_\mu(x, z, p), \\ g^{ab}_{\mu\nu} &= -(w^a_\mu w^b_\nu - w^a_\nu w^b_\mu)/G = (v^a_\mu v^b_\nu - v^a_\nu v^b_\mu)/K. \end{aligned} \tag{5,35}$$

The last equality can be verified immediately by using the expression (5,26c) for w^a_μ. The quantities $g^{ab}_{\mu\nu}$ and $h^{\mu\nu}_{ab} = (\pi^\mu_a \pi^\nu_b - \pi^\nu_a \pi^\mu_b)/L = -(p^\mu_a p^\nu_b - p^\mu_b p^\nu_a)/H_c$ connect the 4 quantities v^a_μ, π^μ_a, p^μ_a and w^a_μ in a symmetrical way: in addition to the relations (5,35) we have

$$\begin{aligned} p^\mu_a &= \pi^\mu_a - h^{\mu\nu}_{ab} v^b_\nu, \qquad \pi^\mu_a = K(p^\mu_a + h^{\mu\nu}_{ab} w^b_\nu)/H_c, \\ w^a_\mu &= v^a_\mu - g^{ab}_{\mu\nu} \pi^\nu_b. \end{aligned} \tag{5,36}$$

Furthermore

$$g^{ab}_{\mu\nu} h^{\mu\nu}_{ab} = 4(K - H_c)/K\,. \tag{5,37}$$

We next want to calculate the value of the functional determinant $(\partial p^{\mu}_{a}/\partial v^{b}_{\nu})$ of the Legendre transform $v \to p$. In order to do this it is convenient to introduce the quantities $k^{\mu}_{a} = p^{\mu}_{a}/H_c$ – compare section 2.5 –. It follows from eq. (5,8) that

$$T^{\mu}_{\rho}\, \mathrm{d}k^{\rho}_{a} = -k^{\rho}_{a}\, \mathrm{d}T^{\mu}_{\rho} + \mathrm{d}\pi^{\mu}_{a}$$
$$= -\,k^{\rho}_{a}(v^{c}_{\rho}\, \mathrm{d}\pi^{\mu}_{c} + \pi^{\mu}_{b}\, \mathrm{d}v^{b}_{\rho} - \delta^{\mu}_{\rho}\, \mathrm{d}L) + \mathrm{d}\pi^{\mu}_{a}$$

and since

$$p^{\rho}_{a} = H_c k^{\rho}_{a} = \pi^{\rho}_{c} Q^{c}_{a}/L\,, \qquad \mathrm{d}L = \pi^{\nu}_{b}\, \mathrm{d}v^{b}_{\nu}\,,$$

we get

$$-T^{\mu}_{\rho}\, \mathrm{d}k^{\rho}_{a} = \frac{1}{H_c} Q^{s}_{a}\Big[\mathrm{d}\pi^{\mu}_{s} - \frac{1}{L}(\pi^{\mu}_{s}\pi^{\rho}_{b} - \pi^{\mu}_{b}\pi^{\rho}_{s})\, \mathrm{d}v^{b}_{\rho}\Big]\,,$$

or, because of eq. (5,12),

$$\mathrm{d}k^{\mu}_{a} = -L^{-1}H_c^{-2} S^{\mu}_{\nu} Q^{s}_{a}\Big[\mathrm{d}\pi^{\nu}_{s} - \frac{1}{L}(\pi^{\nu}_{s}\pi^{\rho}_{b} - \pi^{\rho}_{s}\pi^{\nu}_{b})\, \mathrm{d}v^{b}_{\rho}\Big]\,, \tag{5,38}$$

from which we obtain

$$\partial k^{\mu}_{a}/\partial v^{b}_{\rho} = -L^{-1}H_c^{-2} S^{\mu}_{\nu} Q^{s}_{a}\Big[\frac{\partial^2 L}{\partial v^{b}_{\rho}\, \partial v^{s}_{\nu}} - \frac{1}{L}(\pi^{\nu}_{s}\pi^{\rho}_{b} - \pi^{\rho}_{s}\pi^{\nu}_{b})\Big]\,. \tag{5,39}$$

In order to calculate the determinant of the $2n \times 2n$ matrix (5,39) we observe that

$$|(S^{\mu}_{\nu} Q^{c}_{a})| = |S \times Q| = |S|^{n}\, |Q|^{2}\,,$$

where $S \times Q$ is the Kronecker product of the matrices S and Q [e.g. Bellman, 1970, ch. 12]. According to eqs. (5,11) and (5,22) we have

$$|S| = -LH_c\,, \qquad |Q| = (-1)^{n-1} L H_c^{n-1}\,,$$

and therefore get

$$|(\partial k/\partial v)| = (-1)^{n} L^{2-n} H_c^{-(n+2)} D_c\,, \qquad D_c := \left|\left(\frac{\partial^2 L}{\partial v^{a}_{\mu}\, \partial v^{b}_{\nu}} - \frac{1}{L}(\pi^{\mu}_{a}\pi^{\nu}_{b} - \pi^{\nu}_{a}\pi^{\mu}_{b})\right)\right|\,. \tag{5,40}$$

Furthermore, because

$$\frac{\partial k_a^\mu}{\partial p_b^\nu} = \frac{1}{H_c}\left(\delta_\nu^\mu \delta_a^b - \frac{1}{H_c} p_a^\mu w_\nu^b\right)$$

and since $|(\delta_{ik} + a_i b_k)| = 1 + a_j b_j$ [e.g. Bellman, 1970, p. 83, exercises], we get

$$|(\partial k/\partial p)| = -H_c^{-2n}\left(1 - \frac{1}{H_c} p_a^\mu w_\mu^a\right),$$

which, when combined with eqs. (5,27), gives

$$|(\partial k/\partial p)| = -LH_c^{-2n}/K = |(\partial p/\partial k)|^{-1}\,. \tag{5,41}$$

Eqs. (5,40) and (5,41) provide the final result

$$|(\partial p/\partial v)| = |(\partial p/\partial k)|\,|(\partial k/\partial v)| = (-1)^{n-1} K H_c^{n-2} L^{1-n} D_c\,. \tag{5,42}$$

We see that Carathéodory's generalized Legendre transformation is regular iff $KLH_cD_c \neq 0, \infty$.

5.2. Applications and examples

(i) Since the canonical momentum of a field variable plays an important role for the canonical quantization of a classical system in the framework of quantum field theory [e.g. Bogoliubov and Shirkov, 1959; Bjorken and Drell, 1965; Itzykson and Zuber, 1980], it is important that Carathéodory's definition of the canonical momenta p_a^μ is, at least, compatible with the conventional one, π_a^μ, in the case of free fields: Here it is decisive that for *one* real field, i.e. for $n = 1$, the definitions of p^μ and π^μ coincide. Since for free fields the internal degrees of freedom are quantized independently – the creation and annihilation operators for the two spin components of a free electron are uncoupled and can be treated independently. The same holds for the two polarization states of free photons –. In other words: Suppose that the Lagrangian L for n fields z^a can be decomposed into a sum $L = \Sigma_{a=1}^n L_a$, where each L_a depends only on z^a and v_μ^a, but not on any other z^b and v_μ^b, $b \neq a$. Then for each L_a Carathéodory's canonical theory is the same as the conventional one, because each L_a contains only one internal (field) degree of freedom. More generally, we learn here, that any given Lagrangian L should be decomposed into "irreducible" parts, before applying Carathéodory's formalism!
Example:
The Lagrangian for a free Dirac field $\psi(x)$ can be written as

$$L = \frac{\mathrm{i}}{2}(\bar\psi\gamma^\mu\partial_\mu\psi - (\partial_\mu\bar\psi)\,\gamma^\mu\psi) - m\bar\psi\psi\,,$$

$$\bar\psi = \psi^+\gamma^0\,, \qquad \gamma^0 = \begin{pmatrix} E_2 & 0 \\ 0 & -E_2 \end{pmatrix},$$

$$\gamma^j = \gamma^0\alpha_j\,, \qquad \alpha_j = \begin{pmatrix} 0 & \sigma_j \\ \sigma_j & 0 \end{pmatrix}, \qquad j = 1, 2, 3\,,$$

where the σ_j are Pauli's spin matrices.

Because of the "odd" form of the matrices γ^j the above Lagrangian does not decompose into a sum of 4 terms (4 complex terms plus their 4 complex conjugates), each of which contains only one of the 4 (complex) components of ψ. However, a unitary Foldy–Wouthuysen transformation

$$\psi \to e^{iS}\psi, \qquad S = -i\gamma^j p^j \theta, \qquad \theta = \frac{1}{2|p|}\tan^{-1}\left(\frac{|p|}{m}\right), \qquad p^j = -i\partial_j,$$

diagonalizes the operator $-i\alpha_j\partial_j + m\gamma^0$ and gives the required decomposition [Foldy and Wouthuysen, 1950; Bjorken and Drell, 1964, ch. 4].

(ii) $L = \frac{1}{2}\sum_{a=1}^{2} v^a \cdot v^a - V(z)$, $v^a \cdot v^a := g^{\mu\nu} v^a_\mu v^a_\nu$, where the "potential" term $V(z)$ is assumed to be invariant under O(2) transformations

$$z = \begin{pmatrix} z^1 \\ z^2 \end{pmatrix} \to \hat{z} = Cz,\ C \in \mathrm{O}(2).$$

This invariance implies that we can transform the matrix $(R^{ab}) = (v^a \cdot v^b - \delta^{ab}L)$ into diagonal form. Then the eqs. (5,8) and (5,18) take the form

$$\begin{aligned} p^\mu_a(v^1 \cdot v^1 - L) &= H_c \pi^\mu_a, \qquad \pi^\mu_a = g^{\mu\nu} v^a_\nu, \\ p^\mu_2(v^2 \cdot v^2 - L) &= H_c \pi^\mu_2, \end{aligned} \tag{5,43}$$

and

$$\begin{aligned} H_c &= -\frac{1}{L}(v^1 \cdot v^1 - L)(v^2 \cdot v^2 - L) \\ &= \tfrac{1}{2}(v^1 \cdot v^1 + v^2 \cdot v^2) + V(z) - \frac{1}{L}(v^1 \cdot v^1)(v^2 \cdot v^2). \end{aligned} \tag{5,44}$$

It follows from eqs. (5,43) that

$$(p_1 \cdot p_1)(v^1 \cdot v^1 - L)^2 = H_c^2 v^1 \cdot v^1, \qquad (p_2 \cdot p_2)(v^2 \cdot v^2 - L)^2 = H_c^2 v^2 \cdot v^2,$$

which, together with eq. (5,44), imply

$$L^2(p_1 \cdot p_1)(p_2 \cdot p_2) = H_c^2(v^1 \cdot v^1)(v^2 \cdot v^2). \tag{5,45}$$

Eq. (5,45) relates the two invariants

$$\Delta(v) = (v^1 \cdot v^1)(v^2 \cdot v^2) - (v^1 \cdot v^2)^2,$$

$$\Delta(p) = (p_1 \cdot p_1)(p_2 \cdot p_2) - (p_1 \cdot p_2)^2.$$

From eqs. (5,43) we get in addition

$$p_1 \cdot p_1 + p_2 \cdot p_2 = (v^1 \cdot v^1 + v^2 \cdot v^2)(1 - \Delta(v)/L^2) + 4V\Delta(v)/L^2 \,. \tag{5,46}$$

Combining this relation with eqs. (5,44) and (5,45) we obtain

$$H_c = \tfrac{1}{2}(p_1 \cdot p_1 + p_2 \cdot p_2) + V(z)\,(1 - \Delta(p)/H_c^2)\,. \tag{5,47}$$

For $\Delta(p) \neq 0$ eq. (5,47) is a cubic equation for H_c. This is just one example for the algebraic complexities associated with Carathéodory's canonical theory for fields.

(iii) E-Dynamics in one space and one time dimension – see eq. (3,46) –:

$$L = \tfrac{1}{2}(F_{01})^2 - V\,, \qquad F_{01} = \partial_0 A_1 - \partial_1 A_0\,, \qquad V = j_\alpha A^\alpha\,.$$

We have

$$\pi_0^1 = \pi_1^0 = \partial_0 A^1 + \partial_1 A^0 = -E$$

and

$$\begin{aligned} T_0^0 &= -E\,\partial_0 A^1 - L\,, \qquad T_1^0 = -E\,\partial_1 A^1\,, \\ T_0^1 &= -E\,\partial_0 A^0\,, \qquad T_1^1 = -E\,\partial_1 A^0 - L\,, \end{aligned} \tag{5,48}$$

and therefore the eqs. (5,7) take the form

$$\begin{aligned} Lp_0^0 &= E^2\,\partial_1 A^1\,, \qquad Lp_1^0 = -E^2\,\partial_1 A^0 - EL\,, \\ Lp_0^1 &= -E^2\,\partial_0 A^1 - EL\,, \qquad Lp_1^1 = E^2\,\partial_0 A^0\,. \end{aligned} \tag{5,49}$$

For $V = 0$ the transformations (5,49) are the same as those in eqs. (3,47) for $\lambda = 2$, i.e. the Legendre transformation (5,49) is singular for $V = 0$. This can be seen immediately from the value of the determinant D_c – compare eq. (5,40) –, which in our example has the value $2E^6 V/L^4$. Thus, we have the amusing situation that in the present case Carathéodory's Legendre transformation (5,49) is only regular if the interaction term $j_\alpha A^\alpha$ does not vanish!

The eqs. (5,49) imply

$$L(p_0^0 + p_1^1) = E^2(\partial_0 A^0 + \partial_1 A^1)\,, \tag{5,50a}$$

$$Lp_- = 2EV\,, \qquad p_- := p_1^0 + p_0^1 = p_{01} - p_{10}\,, \tag{5,50b}$$

$$L^2|p| = -E^2|T|\,, \qquad |p| = p_0^0 p_1^1 - p_0^1 p_1^0\,, \tag{5,50c}$$

$$|T| = -L(\tfrac{1}{2}E^2 + V) + E^2(\partial_1 A^0\,\partial_0 A^1 - \partial_0 A^0\,\partial_1 A^1)\,.$$

Because $|T| = -H_c L$, the eqs. (5,50b and c) give

$$H_c L = 4V^2|p|/p_-^2 \tag{5,51}$$

and

$$L = -2VH_c/(2H_c - |p|)\,. \tag{5,52}$$

The last two equations combined lead to a quadratic equation for H_c:

$$H_c^2 + 4V|p|H_c/p_-^2 = 2|p|^2\, V/p_-^2\,, \tag{5,53}$$

with the two roots

$$H_c = -2V|p|/p_-^2 \pm (2/p_-^2)|(V|p|)|(1 + p_-^2/2V)^{1/2}\,. \tag{5,54}$$

Since $1 + p_-^2/2V = (\frac{1}{2}E^2 + V)^2/L^2$, the square root in eq. (5,54) is always real. The choice of its sign depends on the relative sign of V and $|p|$ (or $|T|$) and the relative sign of $\frac{1}{2}E^2 - V$ and $\frac{1}{2}E^2 + V$. The two solutions (5,54) coincide if $p_-^2 = -2V$, which implies $H_c = |p|$, $L = -2V$ or $E^2 = -2j_\alpha A^\alpha$. Up to now we have not imposed any gauge condition. Possible choices of the gauge are: According to eq. (5,50a) the Lorentz gauge $\partial_0 A^0 + \partial_1 A^1 = 0$ implies $p_0^0 + p_1^1 = 0$. A more interesting gauge in the context of the model is $\partial_1 A^0\, \partial_0 A^1 - \partial_0 A^0\, \partial_1 A^1 = 0$, because then, according to the above expression for $|T|$, eq. (5,50c), H_c takes the simple form $H_c = \frac{1}{2}E^2 + V$. The vanishing of the functional determinant $|(\partial_\mu A^\alpha)|$ means that the two functions A^0 and A^1 are not independent. This happens in most gauges: The gauges $A^0 = 0$ or $A^1 = 0$ reduce the number of dependent variables to just one, where Carathéodory's canonical theory is the same as that of DeDonder–Weyl.

(iv) An interesting example for an application of Carathéodory's theory to a physical system is the relativistic string [Kastrup and Rinke, 1981]:
Dependent variables are the four coordinates x^α, $\alpha = 0, 1, 2, 3$, of the Minkowski space M^4, independent variables are the two parameters τ^μ, $\mu = 1, 2$, $-\infty < \tau^1 < +\infty$, $0 \le \tau^2 \le 1$. The dynamics of the string, which was already mentioned in section 4.6, can be derived from the Lagrangian

$$L_s = \tfrac{1}{4} v_{\alpha\beta} v^{\alpha\beta}\,, \qquad v^{\alpha\beta} = v_1^\alpha v_2^\beta - v_2^\alpha v_1^\beta\,, \tag{5,55}$$

too [Schild, 1977]. This Lagrangian gives

$$\begin{aligned} \pi_\alpha^1 &= g_{\alpha\beta}[v_1^\beta (v_2)^2 - v_2^\beta (v_1 \cdot v_2)]\,, \\ \pi_\alpha^2 &= g_{\alpha\beta}[v_2^\beta (v_1)^2 - v_1^\beta (v_1 \cdot v_2)]\,, \end{aligned} \tag{5,56}$$

and therefore

$$T_\nu^\mu = \delta_\nu^\mu L_s\,. \tag{5,57}$$

The eqs. (5,56) and (5,57) imply

$$H_c = -L_s \tag{5,58a}$$

$$p_\alpha^\mu = -\pi_\alpha^\mu\,. \tag{5,58b}$$

We still have to express H_c in terms of the momenta p^μ_α: Because

$$\pi^1_\alpha \pi^2_\beta - \pi^2_\alpha \pi^1_\beta = 2v_{\alpha\beta} L_s ,$$

we get

$$p_{\alpha\beta}\, p^{\alpha\beta} = 4v_{\alpha\beta} v^{\alpha\beta} L_s^2 = 16 L_s^3 , \qquad p_{\alpha\beta} = p^1_\alpha p^2_\beta - p^2_\alpha p^1_\beta .$$

If combined with eq. (5,58a) the last equation gives

$$H_c = -(\tfrac{1}{16} p_{\alpha\beta} p^{\alpha\beta})^{1/3} . \tag{5,59}$$

5.3. Canonical field equations

The canonical field equations in Carathéodory's framework are unpleasantly complicated. In special cases they become simpler and we shall see in section 5.4 that they can be cast into a remarkably simple form by a canonical point transformation! Like in chapter 3 we can derive the canonical field equations by means of the 2-forms

$$i(\partial_a)\, d\Omega_c , \qquad i(\partial/\partial p^\mu_a)\, d\Omega_c , \qquad i(\partial_\mu)\, d\Omega_c ,$$

$$\begin{aligned} d\Omega_c &= H_c^{-2}\, dH_c \wedge \theta^1 \wedge \theta^2 - H_c^{-1} (d\theta^1 \wedge \theta^2 - \theta^1 \wedge d\theta^2) \\ &= -dH_c \wedge (H_c\, dx^1 \wedge dx^2 + \tfrac{1}{2} h_{bs}\, dz^b \wedge dz^s)/H_c + dp^\mu_b \wedge dz^b \wedge d\Sigma_\mu - \varepsilon_{\mu\nu} (p^\mu_b/H_c)\, dp^\nu_s \wedge dz^b \wedge dz^s , \end{aligned}$$

which vanish on the extremals $\hat{\Sigma}^{(2)}_0$. In the following we assume that Carathéodory's Legendre transformation $v^a_\mu \to p^\mu_a$, $L \to H_c$ is regular, i.e. that the determinant D_c in eq. (5,40) does not vanish. Then we know from our general discussion in chapter 3, where the determinant (3,33) here is equal to D_c, that we have $\omega^a = dz^a - \hat{\varphi}^a_\mu\, dx^\mu = 0$ on the extremals $z^a = f^a(x)$, with $\hat{\varphi}^a_\mu(x, z, p)$ given by eq. (5,35). According to eqs. (5,34a) and (5,35) we, therefore, can write the first set of the canonical field equations in the following two ways:

$$\partial_\mu f^a(x)\, V^\mu_\nu(x, z, p) = H_c \frac{\partial H_c}{\partial p^\nu_a}(x, z, p) , \tag{5,60a}$$

or

$$\partial_\mu f^a(x) = -H_c\, G(w^a_\mu + g^{ab}_{\mu\nu}\, p^\nu_b)/|V| = \hat{\varphi}^a_\mu(x, z, p) . \tag{5,60b}$$

The eqs. (5,60) may also be derived from the fact that the 2-forms

$$i(\partial/\partial p^\mu_a)\, d\Omega_c = -w^a_\mu (H_c\, dx^1 \wedge dx^2 + \tfrac{1}{2} h_{bs}\, dz^b \wedge dz^s)/H_c + dz^a \wedge d\Sigma_\mu - \varepsilon_{\mu\rho} (p^\rho_b/H_c)\, dz^a \wedge dz^b =: \omega^a_\mu \tag{5,61}$$

vanish on the extremals $z^a = f^a(x)$.

Example:
From the Hamilton-function (5,59) we obtain

$$\partial_{(1)}x^\alpha = \tfrac{1}{4}H_c^{-2}\, g^{\alpha\beta}[p_\beta^2(p^1\cdot p^2) - p_\beta^1(p^2\cdot p^2)]\,, \qquad \partial_{(2)}x^\alpha = \tfrac{1}{4}H_c^{-2}\, g^{\alpha\beta}[p_\beta^1(p^1\cdot p^2) - p_\beta^2(p^1\cdot p^1)]\,. \tag{5,62}$$

The second set† of the canonical field equations can be derived from the vanishing of the 2-forms

$$i(\partial_a)\,\mathrm{d}\Omega_c = -\partial_a H_c(H_c\,\mathrm{d}x^1\wedge \mathrm{d}x^2 + \tfrac{1}{2}h_{bs}\,\mathrm{d}z^b\wedge \mathrm{d}z^s)/H_c + (h_{ab}/H_c)\,\mathrm{d}H_c\wedge \mathrm{d}z^b - \mathrm{d}p_a^\mu \wedge \mathrm{d}\Sigma_\mu + (1/H_c)\,\varepsilon_{\mu\nu}(p_a^\mu\,\mathrm{d}p_b^\nu - p_b^\mu\,\mathrm{d}p_a^\nu)\wedge \mathrm{d}z^b =: \lambda_a \tag{5,63}$$

for the extremals $z^a = z^a(x) = f^a(x)$, $p_a^\mu = p_a^\mu(x) = g_a^\mu(x)$:

$$(K/H_c)\,\partial_a H_c - (1/H_c)\,h_{ab}^{\mu\nu}\frac{\mathrm{d}H_c}{\mathrm{d}x^\mu}\,\partial_\nu f^b + \partial_\mu p_a^\mu + (1/H_c)[p_b^\mu(\partial_\nu f^b\,\partial_\mu p_a^\nu - \partial_\mu f^b\,\partial_\nu p_a^\nu) - p_a^\mu(\partial_\nu f^b\,\partial_\mu p_b^\nu - \partial_\mu f^b\,\partial_\nu p_b^\nu)] = 0\,, \tag{5,64}$$

where the relation (5,37) has been used. In order to "simplify" the expression (5,64) we can use the equations

$$\mathrm{d}H_c/\mathrm{d}x^\mu = (K/H_c)\,\partial_\mu H_c - \partial_\mu f^b\,\partial_\nu p_b^\nu + \partial_\nu f^b\,\partial_\mu p_b^\nu\,, \tag{5,65}$$

which follow from

$$i(\partial_\mu)\,\mathrm{d}\Omega_c = -\partial_\mu H_c(H_c\,\mathrm{d}x^1\wedge \mathrm{d}x^2 + \tfrac{1}{2}h_{bs}\,\mathrm{d}z^b\wedge \mathrm{d}z^s)/H_c + \mathrm{d}H_c\wedge \mathrm{d}\Sigma_\mu - \varepsilon_{\mu\rho}\,\mathrm{d}p_b^\rho\wedge \mathrm{d}z^b =: \omega_\mu = 0\,. \tag{5,66}$$

The eqs. (5,65) are generalizations of the relation $\mathrm{d}H/\mathrm{d}t = \partial_t H$ in mechanics. Using further that $h_{ab}^{\mu\nu}v_\nu^b = \pi_a^\mu - p_a^\mu$, $p_a^\mu v_\nu^a = -\bar{T}_\nu^\mu + \delta_\nu^\mu H_c$ (see eqs. (5,12)), $\partial_\mu f^a(x) = \hat{\varphi}_\mu^a(x, z, p)$ and $\pi_a^\mu = KV_\nu^\mu p_a^\nu/H_c^2$ (see eqs. (5,8) and (5,31a)), the eqs. (5,64) can be rewritten as

$$C_{a\nu}^{b\mu}\,\partial_\mu p_b^\nu = -H_c\partial_a H_c + C_a^\mu \partial_\mu H_c\,, \qquad C_{a\nu}^{b\mu}(x,z,p) = V_\nu^\mu\,\delta_a^b + p_a^\rho w_\rho^b \delta_\nu^\mu - V_\rho^\mu p_a^\rho \hat{\varphi}_\nu^b/H_c\,, \qquad C_a^\mu(x,z,p) = (\delta_\rho^\mu + GV_\rho^\mu/|V|)\,p_a^\rho\,. \tag{5,67}$$

The eqs. (5,67) may also be derived by differentiating the relation $H_c\pi_a^\mu = T_\nu^\mu p_a^\nu$:

$$\pi_a^\mu\frac{\mathrm{d}H_c}{\mathrm{d}x^\mu} + H_c\frac{\mathrm{d}\pi_a^\mu}{\mathrm{d}x^\mu} = \left(\frac{\mathrm{d}}{\mathrm{d}x^\mu}T_\rho^\mu\right)p_a^\rho + T_\rho^\mu\frac{\mathrm{d}p_a^\rho}{\mathrm{d}x^\mu}$$

and using $\mathrm{d}\pi_a^\mu/\mathrm{d}x^\mu = \partial_a L$, $\mathrm{d}T_\rho^\mu/\mathrm{d}x^\mu = -\partial_\rho L$ and the relations (5,26a + b), (5,27) and (5,31a).

† Due to a copying error the second half of eqs. (7) in [Kastrup, 1977] is not correct. I am indebted to Prof. Géheniau, Dr. Biran [1980] and Dr. David from Brussels for pointing this out to me!

For $n = 1$ we have $C^{1\mu}_{1\nu} = \delta^{\mu}_{\nu} H_c$, $C^{\mu}_{1} = 0$ and the eqs. (5,67) reduce to $\partial_\mu p^\mu = -\partial_z H_c$, as they should. For $n > 1$ the last term $C^{\mu}_{a}\partial_\mu H_c$ in eqs. (5,67) does only occur, if the Lagrangian L contains the variables x explicitly – see eq. (5,26a) –, for instance, if L contains an external current. The eqs. (5,67) take a much simpler form if one writes them in terms of the variables $k^{\mu}_{a} = p^{\mu}_{a}/H_c$: Differentiation of $\pi^{\mu}_{a} = T^{\mu}_{\rho} k^{\rho}_{a}$ and essentially the same arguments as above yield

$$V^{\mu}_{\nu} \frac{\mathrm{d}}{\mathrm{d}x^\mu} k^{\nu}_{a} = -\partial_a H_c - (\partial_\nu H_c)\, k^{\nu}_{a} .$$

As to the interpretation of the variables k^{μ}_{a} – which are the ones which were used by Carathéodory himself (instead of the momenta p^{μ}_{a}!) – see section 5.5.

We further remark that

$$\mathrm{d}\Omega_c = \omega^a \wedge \bar{\lambda}_a ,$$

$$\bar{\lambda}_a = -(K/H_c)\, \partial_a H\, \mathrm{d}x^1 \wedge \mathrm{d}x^2 + (1/2H_c)\, h_{ab}\, \mathrm{d}H_c \wedge (\mathrm{d}z^b + \hat{\varphi}^{b}_{\rho}\, \mathrm{d}x^\rho) - \mathrm{d}p^{\mu}_{a} \wedge \mathrm{d}\Sigma_\mu$$
$$+ (1/H_c)\, \varepsilon_{\mu\nu}(p^{\mu}_{a}\, \mathrm{d}p^{\nu}_{b} \wedge \mathrm{d}z^b - \hat{\varphi}^{b}_{\rho}\, p^{\mu}_{b}\, \mathrm{d}p^{\nu}_{a} \wedge \mathrm{d}x^\rho) ,$$

from which we get

$$\omega^{a}_{\mu} = -\omega^b \wedge i(\partial/\partial p^{\mu}_{a})\, \bar{\lambda}_b \in \mathrm{I}[\omega^a] ,$$

$$\lambda_a = \bar{\lambda}_a - \omega^b \wedge i(\partial_a)\, \bar{\lambda}_b \equiv \bar{\lambda}_a \bmod \mathrm{I}[\omega^a] ,$$

$$\omega_\mu = -\hat{\varphi}^{a}_{\mu} \bar{\lambda}_a - \omega^a \wedge i(\partial_\mu)\, \bar{\lambda}_a \equiv 0 \bmod \mathrm{I}[\omega^a, \lambda_a] .$$

This shows that the 2-forms ω^{a}_{μ} and ω_μ from eqs. (5,63 and 66) lie in the ideal generated by the forms ω^a and λ_a!

5.4. Hamilton–Jacoby theory

The basic relation of Carathéodory's HJ theory for fields is

$$\mathrm{d}S^1(x, z) \wedge \mathrm{d}S^2(x, z) = -\frac{1}{H_c}\, \theta^1 \wedge \theta^2 . \tag{5,68}$$

Comparing coefficients of $\mathrm{d}x^1 \wedge \mathrm{d}x^2$, $\mathrm{d}z^a \wedge \mathrm{d}x^\mu$ and $\mathrm{d}z^a \wedge \mathrm{d}z^b$ on both sides of eq. (5,65) gives

$$|(\partial_\mu S^\nu)| + H_c = 0 , \tag{5,69a}$$

$$p^{\mu}_{a} = (\overline{\partial S})^{\mu}_{\rho}\, \partial_a S^\rho =: \psi^{\mu}_{a}(x, z) , \quad \text{or} \quad p^{\rho}_{a}\, \partial_\rho S^\mu = |(\partial_\rho S^\nu)|\, \partial_a S^\mu = -H_c\, \partial_a S^\mu , \tag{5,69b}$$

$$-(p^{1}_{a} p^{2}_{b} - p^{2}_{a} p^{1}_{b})/H_c = \partial_a S^1\, \partial_b S^2 - \partial_a S^2\, \partial_b S^1 . \tag{5,69c}$$

Inserting $p^{\mu}_{a} = \psi^{\mu}_{a}(x, z)$ into H_c in eq. (5,69a) gives the CHJ partial differential equation for the two functions $S^\mu(x, z)$. The eqs. (5,69b) mean that the vectors $w_{(a)} = p^\mu_a \partial_\mu + H_c\partial_a$ are tangent to the wave fronts $S^\mu(x, z) = \sigma^\mu = \text{const.}$, for we have $dS^\mu(w_{(a)}) = p^\rho_a \partial_\rho S^\mu + H_c\partial_a S^\mu = 0$, where the last equality follows from eqs. (5,69b). The n tangent vectors $w_{(a)}$ of the wave fronts and the 2 tangent vectors $e_{(\mu)} = \partial_\mu + v^a_\mu \partial_a$ of the extremals span an $(n+2)$-dimensional vector space at each point (x, z) if $H_c L \neq 0$ at (x, z). In order to prove this, we calculate the determinant $|(e_{(\mu)}, w_{(a)})|$ of the matrix

$$\begin{pmatrix} E_2 & p \\ v & H_c E_n \end{pmatrix}.$$

With the help of eq. (2,40b) we get

$$|(e_{(\mu)}, w_{(a)})| = |(H_c E_n - v\cdot p)| = (-1)^n |Q| = -H_c^{n-1} L ,$$

where the relation (5,22) has been used.

We shall frequently employ the following notation: In HJ theories the quantities v^a_μ, π^μ_a, and p^μ_a become functions of the variables x and z, $v^a_\mu = \varphi^a_\mu(x, z)$, $p^\mu_a = \psi^\mu_a(x, z)$ etc. Instead of introducing always new symbols φ^a_μ etc. for those functions, we shall alternately use the symbols $\tilde v^a_\mu$, $\tilde p^\mu_a$ etc. for the functions $\varphi^a_\mu(x, z)$, $\psi^\mu_a(x, z)$: $\tilde v^a_\mu = \varphi^a_\mu(x, z)$ etc. For instance, $\tilde L$ and $\tilde T^\mu_\nu$ mean the functions

$$\tilde L(x, z) = L(x, z, v = \varphi(x, z)) , \qquad \tilde T^\mu_\nu(x, z) = \tilde\pi^\mu_a(x, z)\tilde v^a_\nu(x, z) - \delta^\mu_\nu \tilde L .$$

If we know the functions $S^\mu(x, z)$, we can calculate $\tilde p^\mu_a = \psi^\mu_a(x, z)$ and $\tilde w^a_\nu(x, z) = (\partial H_c/\partial p^\nu_a)(x, z, p = \psi(x, z))$ and therefore, according to eq. (5,35), the slope functions

$$\varphi^a_\mu(x, z) = -\tilde H_c\, \tilde G(\tilde w^a_\mu - \tilde g^{ab}_{\mu\nu}\psi^\nu_b)/|\tilde V| . \tag{5,70}$$

Again, the integrability conditions

$$\partial_\nu \varphi^a_\mu + \partial_b\varphi^a_\mu \varphi^b_\nu = \partial_\mu \varphi^a_\nu + \partial_b \varphi^a_\nu \varphi^b_\mu \tag{5,71}$$

will not hold in general.

With $\tilde\omega^a = dz^a - \varphi^a_\rho(x, z)\, dx^\rho$ we have

$$\begin{aligned} dS^\mu &= \partial_\rho S^\mu\, dx^\rho + \partial_a S^\mu(\tilde\omega^a + \varphi^a_\rho\, dx^\rho) \\ &= \Delta^\mu_\rho\, dx^\rho + \partial_a S^\mu \tilde\omega^a , \\ \Delta^\mu_\rho &:= \partial_\rho S^\mu + \varphi^a_\rho \partial_a S^\mu . \end{aligned}$$

Comparing coefficients in

$$dS^1 \wedge dS^2 = L^{-1} a^1 \wedge a^2 , \qquad a^\mu = L\, dx^\mu + \pi^\mu_a \omega^a$$

gives

$$L = \tilde L(x, z) = |\Delta| , \qquad \Delta = (\Delta^\mu_\rho) , \tag{5,72a}$$

$$\tilde{\pi}^\mu_a = \chi^\mu_a(x, z) := (\bar{\Delta})^\mu_\rho \partial_a S^\rho = \partial|\Delta|/\partial v^a_\mu , \tag{5,72b}$$

where the relation (5,24) has been used. Because

$$\delta^\mu_\nu|\Delta| = \Delta^\rho_\nu\bar{\Delta}^\mu_\rho = (\partial_\nu S^\rho + \varphi^a_\nu\partial_a S^\rho)\bar{\Delta}^\mu_\rho = \partial_\nu S^\rho\,\bar{\Delta}^\mu_\rho + \chi^\mu_a\varphi^a_\nu ,$$

we have

$$\tilde{T}^\mu_\nu = \chi^\mu_a\varphi^a_\nu - \delta^\mu_\nu|\Delta| = -\partial_\nu S^\rho\bar{\Delta}^\mu_\rho . \tag{5,73}$$

Suppose $z^a = f^a(x)$ is an arbitrary smooth function and $\varphi^a_\mu(x, z(x)) = \partial_\mu z^a(x)$, then we get $\Delta^\mu_\nu = \mathrm{d}S^\mu(x, z(x))/\mathrm{d}x^\nu$, i.e. $|\Delta|$ is the functional determinant of the transformation

$$x^\mu \to \sigma^\mu = S^\mu(x, z(x))$$

and we have

$$\int_G L[x, z(x), \partial z(x)]\,\mathrm{d}x^1\wedge\mathrm{d}x^2 = \int_G |\Delta|\,\mathrm{d}x^1\wedge\mathrm{d}x^2 = \int_{G_\sigma}\mathrm{d}\sigma^1\wedge\mathrm{d}\sigma^2 = \int_{\partial G_\sigma}\sigma^1\,\mathrm{d}\sigma^2 = V(G_\sigma) , \tag{5,74}$$

where $G_\sigma = \sigma(G)$ is the image of G under the map $x \to \sigma$ and $V(G_\sigma)$ is the volume of G_σ in σ-space! Eq. (5,74) is a generalization of Hilbert's independent integral – compare eq. (2,48) –

$$\int_{t_1}^{t_2} L[t, q(t), \dot{q}(t)]\,\mathrm{d}t = \int_{S_1}^{S_2}\mathrm{d}S = S_2 - S_1 \tag{5,75}$$

from mechanics and was Carathéodory's starting point for his definition of a HJ theory for fields: In analogy to the integral (5,75) which depends only on the boundary values of $q(t)$, $\dot{q}(t)$ at t_1 and t_2 via $S(t_1, q(t_i))$, $i = 1, 2$, the volume integral (5,74) of L over G depends only on the boundary $\partial G_\sigma = \sigma(\partial G)$ of $G_\sigma = \sigma(G)$.

Suppose $z^a = f^a(x)$ is a solution of the equations $\partial_\mu z^a = \varphi^a_\mu(x, z)$, where φ^a_μ is given by eq. (5,70), then $f^a(x)$ is a solution of the Euler–Lagrange equations, too. The proof uses the fact that

$$\frac{\mathrm{d}\bar{\Delta}^\mu_\nu}{\mathrm{d}x^\mu}(x, z(x)) = 0 \tag{5,76}$$

for arbitrary functions $z^a(x)$: Eq. (5,76) follows from the relation

$$\mathrm{d}\bar{A}^\mu_\nu = |A|^{-1}(\bar{A}^\mu_\nu\bar{A}^\rho_\sigma - \bar{A}^\mu_\sigma\bar{A}^\rho_\nu)\,\mathrm{d}A^\sigma_\rho , \tag{5,77}$$

which is a consequence of

$$\delta^\mu_\nu\,\mathrm{d}|A| = \mathrm{d}(A^\mu_\rho\bar{A}^\rho_\nu) = \mathrm{d}(\bar{A}^\mu_\rho A^\rho_\nu)$$

and eq. (5,24). If $A^\sigma_\rho = \Delta^\sigma_\rho = \mathrm{d}S^\sigma(x, z(x))/\mathrm{d}x^\rho$ then eq. (5,77) implies

$$\frac{\mathrm{d}\bar{\Delta}^\mu_\nu}{\mathrm{d}x^\mu} = |\Delta|^{-1}(\bar{\Delta}^\mu_\nu\bar{\Delta}^\rho_\sigma - \bar{\Delta}^\mu_\sigma\bar{\Delta}^\rho_\nu)\frac{\mathrm{d}^2S^\sigma}{\mathrm{d}x^\mu\,\mathrm{d}x^\rho} = 0\,.$$

With the help of eq. (5,76) we obtain from (5,72b)

$$\frac{\mathrm{d}\tilde{\pi}^\mu_a}{\mathrm{d}x^\mu} = \bar{\Delta}^\mu_\rho\frac{\mathrm{d}}{\mathrm{d}x^\mu}\partial_a S^\rho\,, \qquad \tilde{\pi}^\mu_a = \chi^\mu_a(x, z(x))\,.$$

Since

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x^\mu}\partial_a S^\rho &= \partial_\mu\partial_a S^\rho + \partial_b\partial_a S^\rho\,\partial_\mu z^b(x)\\ &= \mathrm{D}_a\Delta^\rho_\mu[x, z, v = \varphi(x, z)] - \partial_b S^\rho\partial_a\varphi^b_\mu\,,\end{aligned}$$

where

$$\mathrm{D}_aF(x, z, v = \varphi(x, z)) = \partial_a F + (\partial F/\partial v^b_\mu)\,\partial_a\varphi^b_\mu\,,$$

we have

$$\begin{aligned}\frac{\mathrm{d}\tilde{\pi}^\mu_a}{\mathrm{d}x^\mu} &= \bar{\Delta}^\mu_\rho(\mathrm{D}_a\Delta^\rho_\mu - \partial_b S^\rho\,\partial_a\varphi^b_\mu)\\ &= \mathrm{D}_a|\Delta| - \chi^\rho_b\,\partial_a\varphi^b_\rho\,.\end{aligned} \tag{5,78}$$

On the other hand

$$\mathrm{D}_a|\Delta| = \mathrm{D}_aL(x, z, v = \varphi(x, z)) = \partial_a L + \tilde{\pi}^\mu_b\partial_a\varphi^b_\mu\,,$$

which, combined with eq. (5,78), gives the Euler–Lagrange equations.
Examples:
(i) $L = \frac{1}{2}((v_0)^2 - (v_1)^2) - V(z)$. In this simple case we have

$$p^\mu = \pi^\mu\,, \qquad p^0 = v_0\,, \qquad p^1 = -v_1\,,$$
$$H_c = \tfrac{1}{2}(p^0)^2 - \tfrac{1}{2}(p^1)^2 + V(z)$$

and the CHJ equation is

$$|(\partial_\mu S^\rho)| + H_c(z, p = \psi(x, z)) = 0\,, \tag{5,79a}$$

$$\begin{aligned}\psi^0(x, z) &= \partial_1 S^1\,\partial_z S^0 - \partial_1 S^0\,\partial_z S^1 = \varphi_0(x, z)\,,\\ \psi^1(x, z) &= \partial_0 S^0\,\partial_z S^1 - \partial_0 S^1\,\partial_z S^0 = -\varphi_1(x, z)\,.\end{aligned} \tag{5,79b}$$

The integrability conditions (5,71) here take the form

$$(\partial_0^2 S^1 - \partial_1^2 S^1)\,\partial_z S^0 - (\partial_0^2 S^0 - \partial_1^2 S^0)\,\partial_z S^1 + |(\partial_\mu S^\nu)|\,(\partial_z S^1\,\partial_z^2 S^0 - \partial_z S^0\,\partial_z^2 S^1)$$
$$+\,\partial_0 S^1\,\partial_0\partial_z S^0 - \partial_0 S^0\,\partial_0\partial_z S^1 + \partial_1 S^0\,\partial_1\partial_z S^1 - \partial_1 S^1\,\partial_1\partial_z S^0 = 0\,. \tag{5,80}$$

Eq. (5,80) becomes simpler, if we make the "separating" ansatz $S^\mu(x, z) = h^\mu(x) + W^\mu(z)$:

$$(\partial_0^2 h^1 - \partial_1^2 h^1)\,\partial_z W^0 - (\partial_0^2 h^0 - \partial_1^2 h^0)\,\partial_z W^1 + |(\partial_\mu h^\nu)|\,(\partial_z W^1\,\partial_z^2 W^0 - \partial_z W^0\,\partial_z^2 W^1) = 0\,. \tag{5,81}$$

If the functions $h^\mu(x)$ are linear in x, but $|(\partial_\mu h^\nu)| \neq 0$, then we must have

$$\partial_z W^1\,\partial_z^2 W^0 - \partial_z W^0\,\partial_z^2 W^1 = 0\,.$$

These conditions are fulfilled by the special ansatz

$$S^0 = -\frac{1}{\sqrt{2}}(\mu^2 + k^2)^{-1/4} A x^0 + (\sqrt{2}/A)\,(\mu^2 + k^2)^{1/4}\,W(z)\,,$$
$$S^1 = \frac{1}{\sqrt{2}}(\mu^2 + k^2)^{1/4} A x^1 - (\sqrt{2}/A)(\mu^2 + k^2)^{1/4} k\,W(z)\,, \tag{5,82}$$
$$\mu, k, A = \text{const.}\,,$$

which gives

$$\psi^0(x, z) = \partial_1 S^1\,\partial_z S^0 = \omega\,W'(z)\,, \qquad \omega = +(\mu^2 + k^2)^{1/2}\,,$$
$$\psi^1(x, z) = \partial_0 S^0\,\partial_z S^1 = k\,W'(z)\,. \tag{5,83}$$

Since $|(\partial_\mu S^\nu)| = -A^2/2$, the CHJ equation

$$-\tfrac{1}{2}A^2 + \tfrac{1}{2}(\psi^0)^2 - \tfrac{1}{2}(\psi^1)^2 + V(z) = 0$$

has the solution

$$W(z) = \frac{1}{\mu}\int^z (A^2 - 2V(\bar z))^{1/2}\,\mathrm{d}\bar z\,. \tag{5,84}$$

(ii) Relativistic string (example iv of section 5.2). We have

$$p^1_\alpha = \partial_{(2)} S^2\,\partial_\alpha S^1 - \partial_{(2)} S^1\,\partial_\alpha S^2\,, \qquad p^2_\alpha = \partial_{(1)} S^1\,\partial_\alpha S^2 - \partial_{(1)} S^2\,\partial_\alpha S^1\,,$$
$$p_{\alpha\beta} = |(\partial_{(\mu)} S^\nu)|\,(\partial_\alpha S^1\,\partial_\beta S^2 - \partial_\alpha S^2\,\partial_\beta S^1) =: \psi_{\alpha\beta}(\tau, x)\,,$$

and from eq. (5,59) we get the equation

$$|(\partial_{(\mu)} S^\nu)| - (\tfrac{1}{16}\psi_{\alpha\beta}\psi^{\alpha\beta})^{1/3} = 0\,.$$

It has the special solution [Kastrup and Rinke, 1981]:

$$S^1 = -\tau^1 + 2(1-\omega^2\rho^2)^{-1}(x^0 - \omega\rho^2\theta)\,,$$
$$S^2 = \tfrac{1}{2}\tau^2 - g(\rho)\,, \qquad \mathrm{d}g/\mathrm{d}\rho = (1-\omega^2\rho^2)^{1/2}\,, \tag{5,85}$$
$$\rho = [(x^1)^2 + (x^2)^2]^{1/2}\,, \qquad \theta = \tan^{-1}(x^2/x^1)\,, \qquad \omega = \text{const.}\,.$$

The solution (5,85) "embeds" the string motion

$$x^0 = \tau^1\,, \qquad x^1(\tau) = \rho(\tau^2)\cos\omega\tau^1\,, \qquad x^2(\tau) = \rho(\tau^2)\sin\omega\tau^1\,,$$
$$x^3 = 0\,, \qquad \mathrm{d}\rho/\mathrm{d}\tau^2 = (1-\omega^2\rho^2)^{-1/2}\,, \qquad 0 \le \rho < \omega^{-1}\,, \tag{5,86}$$

which has the momenta

$$p^1_\alpha = g_{\alpha\beta}\,\partial_{(1)}x^\beta(1-\omega^2\rho^2)^{-1}\,,$$
$$p^2_\alpha = -g_{\alpha\beta}\,\partial_{(2)}x^\beta(1-\omega^2\rho^2)\,. \tag{5,87}$$

On the other hand we get from the solution (5,85)

$$\psi^1_\alpha = (1-\omega^2\rho^2)^{-1} f_\alpha(x) + \tfrac{1}{2}h(x)\,\partial_{\alpha\rho}\,,$$
$$(f_\alpha) = (1, \omega x^2, -\omega x^1, 0)\,,$$
$$h(x) = 4\omega\rho(1-\omega^2\rho^2)^{-2}(\omega x^0 - \theta)\,, \tag{5,88}$$
$$\psi^2_\alpha = (1-\omega^2\rho^2)^{1/2}\partial_\alpha\rho\,.$$

These functions ψ^μ_α coincide with the momenta (5,87) on the extremal (5,86). For the slope functions $\varphi^\alpha_\mu(\tau, x)$ we obtain from the functions (5,88) – compare eqs. (5,62) –

$$\varphi^\alpha_1 = g^{\alpha\beta} f_\beta(x)\,,$$
$$\varphi^\alpha_2 = -(1-\omega^2\rho^2)^{-1/2}\, g^{\alpha\beta}[\tfrac{1}{2}h(x)\, f_\beta(x) + \partial_\beta\rho]\,. \tag{5,89}$$

These slope functions obey the integrability conditions (5,71) everywhere in M^4 as long as they are well-defined, not just on the extremals (5,86).

5.5. The geometrical background of Carathéodory's canonical framework

Let me briefly discuss the geometrical origin of Carathéodory's canonical theory. It is intimately related to our discussion of contact transformations in chapter 2. Because it makes hardly any difference, we assume in this section that the number of independent variables is $m > 1$, not just 2.

A system of n-dimensional wave fronts in $G^{n+m} \subset R^{n+m}$ can be given by the m equations

$$S^\mu(x, z) = \sigma^\mu = \text{const.}\,, \qquad \mu = 1, \ldots, m\,. \tag{5,90}$$

Let $C = \{x(\tau), z(\tau)\}$ be a curve inside the wave front defined by eqs. (5,90). Any tangent vector with components $\dot{x}^\mu$, $\dot{z}^a$ of such a curve has to obey the equations

$$\partial_\nu S^\mu \dot{x}^\nu + \partial_a S^\mu \dot{z}^a = 0 .$$

If we take the special curves $\tau = z^a$, $a = 1, \ldots, n$, we get

$$\partial_\nu S^\mu \, k^\nu_a + \partial_a S^\mu = 0 , \qquad k^\nu_a = \partial_a x^\nu . \tag{5,91}$$

If we define the differential operator

$$\mathrm{d}/\mathrm{d}z^a := \partial_a + k^\mu_a \partial_\mu ,$$

the eqs. (5,91) can be written as $\mathrm{d}S^\mu/\mathrm{d}z^a = 0$. Since the quantities k^μ_a are derivatives with respect to z^a, see eq. (5,91), they satisfy the relations

$$\frac{\mathrm{d}}{\mathrm{d}z^a} k^\mu_b = \frac{\mathrm{d}}{\mathrm{d}z^b} k^\mu_a . \tag{5,92}$$

Furthermore, because of eqs. (5,91) we have

$$\frac{\mathrm{d}}{\mathrm{d}z^a} \partial_\mu S^\nu = -\partial_\rho S^\nu \, \partial_\mu k^\rho_a ,$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}z^a} |(\partial_\mu S^\nu)| = -|(\partial_\mu S^\nu)| \, \partial_\rho k^\rho_a . \tag{5,93}$$

If we define

$$p^\mu_a = -|(\partial_\rho S^\nu)| \, k^\mu_a ,$$

then, according to eqs. (5,91), these p^μ_a obey eqs. (5,69b) and we see the geometrical background for Carathéodory's definition of the canonical momenta k^μ_a (or p^μ_a): By starting with the wave fronts the geometrically natural generalization of the canonical momenta in mechanics – see chapter 2 – appears to be Carathéodory's choice, which is, however, algebraically much more complicated than the conventional one, π^μ_a.

The Euler–Lagrange equations for m-dimensional surfaces $z^a = f^a(x)$ which "traverse" the wave fronts can be obtained as follows: On the intersection the functions $S^\mu(x, z)$ become $S^\mu(x, z = f(x))$. Abbreviating the derivatives $\partial_\mu f^a(x)$ by v^a_μ we have

$$\frac{\mathrm{d}}{\mathrm{d}x^\nu} S^\mu(x, z(x)) = \partial_\nu S^\mu + v^a_\nu \partial_a S^\mu =: \Delta^\mu_\nu$$

and "define", see eqs. (5,72a and b),

$$\pi^{\mu}_{a} = \partial|\Delta|/\partial v^{a}_{\mu} = (\bar{\Delta})^{\mu}_{\rho}\, \partial_a S^{\rho} \tag{5,94a}$$

and

$$T^{\mu}_{\nu} = \pi^{\mu}_{a} v^{a}_{\nu} - \delta^{\mu}_{\nu}|\Delta| = -\partial_{\nu} S^{\rho} \bar{\Delta}^{\mu}_{\rho}\,. \tag{5,94b}$$

Because $|(\bar{\Delta}^{\mu}_{\nu})| = |\Delta|^{m-1}$, the last equation implies

$$|T| = -(-|\Delta|)^{m-1}\, |(\partial_{\mu} S^{\nu})|\,. \tag{5,95}$$

If we define $H_c := |T|(-|\Delta|)^{1-m}$, eq. (5,95) means $H_c + |(\partial_{\mu} S^{\nu})| = 0$. Since

$$k^{\mu}_{a} = -(\overline{\partial S})^{\mu}_{\rho}\, \partial_a S^{\rho}/|(\partial_{\lambda} S^{\nu})|\,,$$

we get from eqs. (5,94a and b):

$$T^{\mu}_{\rho} k^{\rho}_{a} = \pi^{\mu}_{a}\,, \qquad \text{or } T^{\mu}_{\rho} p^{\rho}_{a} = -|(\partial_{\lambda} S^{\nu})|\pi^{\mu}_{a}\,. \tag{5,96}$$

Introducing the matrix

$$S^{\mu}_{\nu} = p^{\mu}_{a} v^{a}_{\nu} - \delta^{\mu}_{\nu} H_c = (\overline{\partial S})^{\mu}_{\rho}\, \Delta^{\rho}_{\nu}\,, \tag{5,97}$$

we have

$$T^{\mu}_{\rho} S^{\rho}_{\nu} = -\delta^{\mu}_{\nu}|(\partial_{\lambda} S^{\rho})|\, |\Delta| = \delta^{\mu}_{\nu} H_c |\Delta|\,, \tag{5,98}$$

from which we obtain

$$|T|\, |S| = H_c^m |\Delta|^m\,,$$

or, because of eq. (5,95),

$$|S| = (-H_c)^{m-1}\, |\Delta|\,. \tag{5,99}$$

The eqs. (5,98) and (5,99) imply that

$$\bar{S}^{\mu}_{\nu} = -(-H_c)^{m-2} T^{\mu}_{\nu}\,. \tag{5,100}$$

Taking the differential of the relation (5,99) we get

$$-(m-1)(-H_c)^{m-2}\, |\Delta|\, dH_c + (-H_c)^{m-1}\, d|\Delta| = \bar{S}^{\mu}_{\nu}\, dS^{\nu}_{\mu} = -(-H_c)^{m-2}\, T^{\mu}_{\nu}\, d(p^{\nu}_{a} v^{a}_{\mu} - \delta^{\nu}_{\mu} H_c)\,,$$

and therefore

$$K\, dH_c - T^{\nu}_{\mu}\, v^{a}_{\nu}\, dp^{\mu}_{a} = -H_c\, (d|\Delta| - \pi^{\mu}_{a}\, dv^{a}_{\mu})\,, \qquad K = v^{a}_{\mu}\pi^{\mu}_{a} - |\Delta|\,, \tag{5,101}$$

from which we get

$$w^a_\mu = \partial H_c/\partial p^\mu_a = v^a_\nu T^\nu_\mu/K\,, \tag{5,102a}$$

$$K\,\partial_a H_c = -H_c\,\partial_a|\Delta|\,, \tag{5,102b}$$

$$K\,\partial_\mu H_c = -H_c\,\partial_\mu|\Delta|\,. \tag{5,102c}$$

The eq. (5,101) shows that H_c is to be considered as a function of x, z, p, if $|\Delta|$ is a function of x, z and v.
In complete analogy to the discussion following the eqs. (5,26) we can derive the relations

$$T^\mu_\nu/K = V^\mu_\nu/H_c\,, \qquad V^\mu_\nu = p^\mu_a\,w^a_\nu - \delta^\mu_\nu G\,,$$
$$v^a_\nu V^\nu_\mu = w^a_\mu H_c\,, \qquad G = p^\mu_a w^a_\mu - H_c\,. \tag{5,103}$$

Taking the differential of $T^\mu_\rho k^\rho_a = \pi^\mu_a$ we get, with the help of the matrix

$$(\tilde{Q}^a_b) = (v^a_\mu k^\mu_b - \delta^a_b)\,, \qquad \tilde{Q}^b_a \pi^\mu_b = L k^\mu_a\,, \qquad L = |\Delta|\,,$$

the relation

$$-T^\mu_\rho\,\mathrm{d}k^\rho_a = \tilde{Q}^b_a\left[\mathrm{d}\pi^\mu_b - \frac{1}{L}(\pi^\mu_b\,\mathrm{d}|\Delta| - \pi^\mu_c\pi^\nu_b\,\mathrm{d}v^c_\nu)\right], \tag{5,104}$$

from which we obtain, with $z^a = f^a(x)$,

$$-T^\mu_\rho\,\mathrm{d}k^\rho_a/\mathrm{d}x^\mu + k^\mu_a\,\partial_\mu|\Delta| + \partial_a|\Delta| = \tilde{Q}^b_a\left[\frac{\mathrm{d}\pi^\mu_b}{\mathrm{d}x^\mu} - \partial_b|\Delta| - \frac{1}{L}(\pi^\mu_b\pi^\rho_c - \pi^\mu_c\pi^\rho_b)\frac{\mathrm{d}v^c_\rho}{\mathrm{d}x^\mu}\right]. \tag{5,105}$$

With the help of eqs. (5,102) and (5,103) the l.h. side of the last equation becomes

$$-Kw^b_\rho(\partial_b k^\rho_a + k^\mu_b\,\partial_\mu k^\rho_a) + (KG/H_c)\,\partial_\mu k^\mu_a - (K/H_c)(\partial_a H_c + k^\mu_a\,\partial_\mu H_c)\,.$$

Because of the definition of $\mathrm{d}/\mathrm{d}z^a$ and since

$$\frac{\mathrm{d}H_c}{\mathrm{d}z^a} = w^b_\nu\frac{\mathrm{d}p^\nu_b}{\mathrm{d}z^a} + \partial_a H_c + k^\mu_a\partial_\mu H_c\,, \qquad \frac{\mathrm{d}p^\nu_b}{\mathrm{d}z^a} = \frac{\mathrm{d}H_c}{\mathrm{d}z^a}k^\nu_b + H_c\frac{\mathrm{d}k^\nu_b}{\mathrm{d}z^a}\,,$$

the relation (5,105) finally takes the form

$$-Kw^b_\rho\left(\frac{\mathrm{d}k^\rho_a}{\mathrm{d}z^b} - \frac{\mathrm{d}k^\rho_b}{\mathrm{d}z^a}\right) + \frac{L}{H_c}\left(\frac{\mathrm{d}H_c}{\mathrm{d}z^a} + H_c\,\partial_\mu k^\mu_a\right) = \tilde{Q}^b_a\left[\frac{\mathrm{d}\pi^\mu_b}{\mathrm{d}x^\mu} - \partial_b L - \frac{1}{L}\,(\pi^\mu_b\pi^\nu_c - \pi^\mu_c\pi^\nu_b)\frac{\mathrm{d}v^c_\nu}{\mathrm{d}x^\mu}\right], \tag{5,106}$$

where the properties $G/H_c + 1 = k^\mu_a w^a_\mu$, $GK = H_c L$ have been used. The l.h. side of eqs. (5,106) vanish, because the equalities (5,92) and (5,93) hold on the wave fronts. On the r.h. side the last term vanishes because $\mathrm{d}v^c_\nu/\mathrm{d}x^\mu = \partial^2 f^c/\partial x^\nu\partial x^\mu$ is symmetric in ν and μ and we are left with the homogeneous equations

$$\tilde{Q}^b_a(\mathrm{d}\pi^\mu_b/\mathrm{d}x^\mu - \partial_b L) = 0\,. \tag{5,107}$$

Since $|\tilde{Q}| = (-1)^{n-1}L/H_c$, the eqs. (5,107) imply the Euler–Lagrange equations

$$\mathrm{d}\pi^\mu_a/\mathrm{d}x^\mu - \partial_a L = 0\,,$$

if $L/H_c \neq 0, \infty$.

As the Euler–Lagrange equations are "dynamical" equations, one wonders which is the crucial geometrical input which determines the dynamics. It seems to me that the identification (5,94a) is the crucial one, together, of course, with the assumption that the determinant $|\Delta|$ is to be identified with the Lagrangian L! In addition we have the important consistency requirement that the derivatives $\partial_\mu f^a(x)$ have to be equal to the slope functions (5,70).

5.6. CHJ currents and complete integrals

We next discuss the conserved current associated with a solution $S^\mu(x, z; a)$ of the CHJ equation which depends on a parameter a. As to the notation, it is useful to introduce the following convention: Let $z^b = f^b(x)$ be an arbitrary smooth function, not necessarily an extremal. Then we shall denote the functions

$$p^\mu_b(x; a) = \psi^\mu_b(x, z = f(x); a) = \overline{(\partial S)}{}^\mu_\rho\, \partial_b S^\rho[x, z = f(x); a]\,,$$
$$H_c(x, z = f(x), p(x; a)) \text{ etc.}$$

by $\mathring{p}^\mu_a$ and $\mathring{H}_c$ etc.

It follows from

$$\mathrm{d}\mathring{S}^\mu = \partial_\rho S^\mu\, \mathrm{d}x^\rho + \partial_\rho z^b(x)\partial_b S^\mu\, \mathrm{d}x^\rho + \frac{\partial S^\mu}{\partial a}\mathrm{d}a$$
$$= \mathring{\Delta}^\mu_\rho\, \mathrm{d}x^\rho + \frac{\partial S^\mu}{\partial a}\mathrm{d}a$$

that

$$\mathrm{d}\mathring{S}^1 \wedge \mathrm{d}\mathring{S}^2 = -\mathring{H}_c^{-1}\mathring{\theta}^1 \wedge \mathring{\theta}^2 + G^\mu\, \mathrm{d}a \wedge \mathrm{d}\Sigma_\mu\,, \tag{5,108}$$

where

$$\mathring{\theta}^\mu = -\mathring{H}_c\, \mathrm{d}x^\mu + \mathring{p}^\mu_b\, \mathrm{d}z^b(x)$$
$$= (\mathring{p}^\mu_b\, \partial_\rho z^b - \delta^\mu_\rho \mathring{H}_c)\, \mathrm{d}x^\rho = \mathring{S}^\mu_\rho\, \mathrm{d}x^\rho$$

and

$$\tilde{G}^{\mu}(x, z; \alpha) = \frac{\partial}{\partial a} S^{\rho}(x, z; a)\, \bar{\Delta}^{\mu}_{\rho}(x, z; a)\,, \qquad (5,109)$$
$$G^{\mu}(x; a) = \tilde{G}^{\mu}(x, z = f(x); a)\,.$$

Since

$$-\mathrm{d}(\mathring{H}_c^{-1}\mathring{\theta}^1 \wedge \mathring{\theta}^2) = \mathring{H}_c^{-2}\, \mathrm{d}\mathring{H}_c \wedge \mathring{\theta}^1 \wedge \mathring{\theta}^2 + \mathring{H}_c^{-1}(-\mathrm{d}\mathring{\theta}^1 \wedge \mathring{\theta}^2 + \mathring{\theta}^1 \wedge \mathrm{d}\mathring{\theta}^2)\,,$$
$$\mathrm{d}\mathring{H}_c = \frac{\mathrm{d}\mathring{H}_c}{\mathrm{d}x^{\rho}}\mathrm{d}x^{\rho} + \frac{\partial H_c}{\partial p^{\rho}_b}\frac{\partial \mathring{p}^{\rho}_b}{\partial a}\mathrm{d}a\,, \qquad \mathring{\theta}^1 \wedge \mathring{\theta}^2 = |\mathring{S}|\, \mathrm{d}x^1 \wedge \mathrm{d}x^2\,,$$
$$\mathrm{d}\mathring{\theta}^{\mu} = -\mathrm{d}\mathring{H}_c \wedge \mathrm{d}x^{\mu} + \partial_{\rho} z^b\, \mathrm{d}\mathring{p}^{\mu}_b \wedge \mathrm{d}x^{\rho}\,,$$
$$\mathrm{d}\mathring{p}^{\mu}_b = \frac{\mathrm{d}\mathring{p}^{\mu}_b}{\mathrm{d}x^{\rho}}\mathrm{d}x^{\rho} + \frac{\partial \mathring{p}^{\mu}_b}{\partial a}\mathrm{d}a\,,$$

we get

$$-\mathrm{d}(\mathring{H}_c^{-1}\mathring{\theta}^1 \wedge \mathring{\theta}^2) = [(\mathring{H}_c^{-2}|\mathring{S}| + \mathring{H}_c^{-1}\mathring{S}^{\mu}_{\mu})\mathring{w}^b_{\rho} + \mathring{H}_c^{-1}\mathring{T}^{\mu}_{\rho}\, \partial_{\mu} z^b]\frac{\partial \mathring{p}^{\rho}_b}{\partial a}\mathrm{d}a \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2\,,$$

where the relation $\bar{S}^{\mu}_{\nu} = -T^{\mu}_{\nu}$ has been used. According to eq. (5,100) we have for $m = 2$: $|\mathring{S}| = -\mathring{H}_c|\mathring{\Delta}|$, and since

$$\mathring{S}^{\mu}_{\mu} = -\mathring{T}^{\mu}_{\mu} = 2|\mathring{\Delta}| - \mathring{\pi}^{\mu}_a\, \partial_{\mu} z^a = |\mathring{\Delta}| - \mathring{K}\,,$$

we finally get

$$-\mathrm{d}(\mathring{H}_c^{-1}\mathring{\theta}^1 \wedge \mathring{\theta}^2) = \mathring{H}_c^{-1}(\partial_{\mu} z^b \mathring{T}^{\mu}_{\rho} - \mathring{K}\mathring{w}^b_{\rho})\frac{\partial \mathring{p}^{\rho}_b}{\partial a}\mathrm{d}a \wedge \mathrm{d}x^1 \wedge \mathrm{d}x^2\,. \qquad (5,110)$$

Taking the exterior derivative of eq. (5,108) we therefore obtain the basic relation

$$\mathring{H}_c^{-1}(\partial_{\mu} z^b \mathring{T}^{\mu}_{\rho} - \mathring{K}\mathring{w}^b_{\rho})\frac{\partial \mathring{p}^{\rho}_b}{\partial a} + \frac{\mathrm{d}G^{\mu}}{\mathrm{d}x^{\mu}} = 0\,. \qquad (5,111)$$

We now can argue in the same way as in chapter 4:
1. If $z^a = f^a(x)$ is an extremal, it follows from eq. (5,26c) that

$$\partial_{\mu} z^b \mathring{T}^{\mu}_{\rho} - \mathring{K}\mathring{w}^b_{\rho} = 0$$

and we see that the current (5,109) is conserved "along" the extremal $z^a = f^a(x)$ which is weakly embedded in the geodesic field $S^{\mu}(x, z; a)$.
Examples:
(i) The solution (5,82) depends on the 2 parameters A and k. The current (5,109) associated with the parameter A is

$$\tilde{G}^0_A = -\tfrac{1}{2}Ax^0 + A^{-1}[x^0 - \omega(\omega x^0 - kx^1)](A^2 - 2V(z)) - A^{-1}\omega\, W(z) + A\omega \int^z (A^2 - 2V(\bar{z}))^{-1/2}\, d\bar{z}\,, \tag{5,112}$$

$$\tilde{G}^1_A = -\tfrac{1}{2}Ax^1 + A^{-1}[x^1 - k(\omega x^0 - kx^1)]\,(A^2 - 2V(z)) - A^{-1}k\, W(z) + Ak \int^z (A^2 - 2V(\bar{z}))^{-1/2}\, d\bar{z}\,,$$

with $\omega^2 = k^2 + 1$, $\mu = 1$. For $V(z) = \frac{1}{2}z^2$ components (5,112) take the form

$$G^0_A = -\tfrac{1}{2}Ax^0 + A^{-1}[x^0 - \omega(\omega x^0 - kx^1)]\,(A^2 - z^2) - \tfrac{1}{2}A^{-1}\omega z(A^2 - z^2)^{1/2} + \tfrac{1}{2}A\omega \sin^{-1}(z/A)\,,$$

$$G^1_A = -\tfrac{1}{2}Ax^1 + A^{-1}[x^1 - k(\omega x^0 - kx^1)]\,(A^2 - z^2) - \tfrac{1}{2}A^{-1}kz(A^2 - z^2)^{1/2} + \tfrac{1}{2}Ak \sin^{-1}(z/A)\,,$$

$$z(x) = A \sin(\omega x^0 - kx^1)\,.$$

The current $(\tilde{G}^\mu_k)$ associated with the parameter k is

$$\tilde{G}^0_k = \tfrac{1}{4}A^2k\omega^{-2}x^0 + \tfrac{1}{2}k^2\omega^{-2}(kx^0 + \omega x^1)(A^2 - V(z)) + \tfrac{1}{2}k\omega^{-1}\, W(z) - 2A^{-2}k\omega W(z)\,(A^2 - 2V(z))\,,$$

$$\tilde{G}^1_k = -\tfrac{1}{4}A^2k\omega^{-2}x^1 + \tfrac{1}{2}k\omega^{-1}(kx^0 + \omega x^1)(A^2 - V(z)) + (1 + \tfrac{1}{2}k^2\omega^{-2})\, W(z) - 2A^{-2}\omega^2\, W(z) \times (A^2 - 2V(z))\,, \tag{5,113}$$

$$W(z) = \int^z (A^2 - 2V(\bar{z}))^{1/2}\, d\bar{z}\,.$$

As in the last chapter I stress the fact that the currents (5,112) and (5,113) are not conserved for arbitrary solutions of the field equations, but only for those which are embedded in the corresponding geodesic field $S^\mu(x, z; a)$. In addition we see that the currents (5,112) and (5,113) do not have a well-defined global charge

$$\int_{\mathbf{R}} G^0_{A,k}\, dx^1\,,$$

because these integrals diverge.

(ii) Noether's theorem, once more.

Let $\{g(a)\}$ be a 1-parameter transformation group

$$x^\mu \to [g(a)(x)]^\mu = \hat{x}^\mu(x, z; a)\,,$$
$$\hat{x}^\mu(x, z; a = 0) = x^\mu\,,$$
$$z^b \to [g(a)(z)]^b = \hat{z}^b(x, z; a)\,,$$
$$\hat{z}^b(x, z; a = 0) = z^b\,,$$

such that

$$\mathrm{d}S^1(\hat{x}(a), \hat{z}(a)) \wedge \mathrm{d}S^2(\hat{x}(a), \hat{z}(a)) = \mathrm{d}S^1(x, z) \wedge \mathrm{d}S^2(x, z)\,.$$

With the definitions

$$\frac{\partial \hat{x}^\mu}{\partial a}\bigg|_{a=0} = X^\mu(x, z)\,, \qquad \frac{\partial \hat{z}^b}{\partial a}\bigg|_{a=0} = Z^b(x, z)$$

we have

$$\frac{\partial S^\rho}{\partial a}[\hat{x}(a), \hat{z}(a)]|_{a=0} = \partial_\nu S^\rho X^\nu + \partial_b S^\rho Z^b$$

and the current (5,109) becomes

$$\tilde{G}^\mu = \frac{\partial S^\rho}{\partial a}\bigg|_{a=0} \bar{\Delta}^\mu_\rho = (\partial_\nu S^\rho \bar{\Delta}^\mu_\rho) X^\nu + (\partial_b S^\rho \bar{\Delta}^\mu_\rho) Z^b = -\tilde{T}^\mu_\nu X^\nu + \tilde{\pi}^\mu_b Z^b\,, \tag{5,114}$$

where the relations (5,72b) and (5,73) have been used. For an extremal $z^a = f^a(x)$ the expression (5,114) becomes the usual Noether current associated with an invariance group $g(a)$ of the Lagrangian form $L\,\mathrm{d}x^1 \wedge \mathrm{d}x^2$ [Noether, 1918; Hill, 1951].

2. Complete integral. As in the case of the DWHJ theory we can define a complete integral for the CHJ theory, too: Let $S^\mu(x, z; a)$ be a solution of the CHJ equation which depends on $2n$ parameters a^ν_b, $\nu = 1, 2$, $b = 1, \ldots, n$ such that

$$|(\partial \psi^\mu_b / \partial a^\nu_c)| \neq 0\,, \tag{5,115}$$

then $S^\mu(x, z; a)$ is called a "complete integral" and the $2n$ currents

$$G^{\mu;b}_\nu(x) = \frac{\partial S^\rho}{\partial a^\nu_b} \bar{\Delta}^\mu_\rho(x, z = f(x))$$

are conserved along any extremal $z^a = f^a(x)$ which is embedded in the geodesic field $S^\mu(x, z; a)$. As in chapter 4 let us assume we have found $4n$ functions $g^{\mu;b}_\nu(x)$ of x which obey the equations

$$\frac{\mathrm{d}}{\mathrm{d}x^\mu} g^{\mu;b}_\nu(x) = 0 \tag{5,116}$$

identically in x for all ν and b and for which the $4n$ equations

$$\tilde{G}^{\mu;b}_\nu(x, z; a) = g^{\mu;b}_\nu(x) \tag{5,117}$$

have n solutions $z^a = f^a(x)$, then the functions $f^a(x)$ are extremals. The proof is completely analogue to that in the case of the DWHJ theory: Inserting the expressions (5,117) into eq. (5,111) it follows then from the property (5,116) that

$$(\partial_\rho z^b(x) \mathring{T}^\rho_\mu - \mathring{K}\mathring{w}^b_\mu) \partial \mathring{p}^\mu_b / \partial a^\nu_c = 0\,, \tag{5,118}$$

which, due to the inequality (5,115), implies

$$\partial_\rho z^b(x)\mathring{T}^\rho_\mu - \mathring{K}\mathring{w}^b_\mu = 0$$

and we see that the functions $f^b(x)$ obey the first set (5,60) of the canonical equations. Furthermore, we have proven above that any such function, embedded in a field $S^\mu(x, z; a)$ obeys the Euler–Lagrange equations which, on the other hand, guarantees the validity of the second half (5,67) of the canonical field equations!

Example:

For the solution (5,82) we obtain

$$\left|\begin{pmatrix} \partial\psi^0/\partial A & \partial\psi^0/\partial k \\ \partial\psi^1/\partial A & \partial\psi^1/\partial k \end{pmatrix}\right| = A/\omega \neq 0\,,$$

which shows that it constitutes a complete integral.

5.7. *The canonical E. Hölder transformation*

In mechanics a transformation

$$t \to \hat{t} = \hat{T}(t, q, p)\,, \qquad q^j \to \hat{q}^j = \hat{Q}^j(t, q, p)\,, \qquad p_j \to \hat{p}_j = \hat{P}_j(t, q, p)$$

is conventionally called canonical* if there exist 2 functions $\hat{H}$ and F such that the following relation holds

$$\begin{aligned} \hat{\theta} &:= -\hat{H}\,\mathrm{d}\hat{t} + \hat{p}_j\,\mathrm{d}\hat{q}^j \\ &= \theta + \mathrm{d}F = -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j + \mathrm{d}F\,. \end{aligned} \qquad (5,119)$$

Since $\mathrm{d}\hat{\theta} = \mathrm{d}\theta$ the equations of motion derived from $\hat{\theta}$ have the same form as those of θ:

$$\mathrm{d}\hat{q}^j/\mathrm{d}\hat{t} = \partial\hat{H}/\partial\hat{p}_j\,, \qquad \mathrm{d}\hat{p}_j/\mathrm{d}\hat{t} = -\partial\hat{H}/\partial\hat{q}^j\,,$$

The function F which can depend on any of the sets of variables (t, q, p), $(\hat{t}, \hat{q}, \hat{p})$, $(t, q, \hat{p})$, $(t, \hat{q}, p)$ etc. is called the "generating function" of the canonical transformation. It follows, for instance, from eq. (2,14) that the time evolution $(t_0, q(t_0), p(t_0)) \to (t, q(t), p(t))$ of a physical system is a canonical transformation which is generated by the function $S(t, q; t_0, q_0)$. The situation becomes more complicated if we try to define canonical transformations for a field theory characterized by the form

$$\Omega = -H\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \varepsilon_{\mu\nu}p^\mu_a\,\mathrm{d}z^a \wedge \mathrm{d}x^\nu + \eta_{ab}\,\mathrm{d}z^a \wedge \mathrm{d}z^b\,.$$

At first sight a transformation

$$\begin{aligned} &x^\mu \to \hat{x}^\mu = \hat{X}^\mu(x, z, p)\,, \qquad z^a \to \hat{z}^a = \hat{Z}^a(x, z, p)\,, \\ &p^\mu_a \to \hat{p}^\mu_a = \hat{P}^\mu_a(x, z, p)\,, \end{aligned} \qquad (5,120)$$

*See however Abraham and Marsden [1978] and Arnold [1978] who use a different definition.

might be called "canonical", if there exist functions $\hat{H}$ and $\hat{\eta}_{ab}$ and an exact 2-form $\mathrm{d}\Sigma$, such that

$$\hat{\Omega} = \Omega + \mathrm{d}\Sigma\,,$$
$$\hat{\Omega} = -\hat{H}\,\mathrm{d}\hat{x}^1 \wedge \mathrm{d}\hat{x}^2 + \varepsilon_{\mu\nu}\hat{p}^{\mu}_{a}\,\mathrm{d}\hat{z}^a \wedge \mathrm{d}\hat{x}^{\nu} + \hat{\eta}_{ab}\,\mathrm{d}\hat{z}^a \wedge \mathrm{d}\hat{z}^b$$

has the same rank as Ω. Since $\mathrm{d}\hat{\Omega} = \mathrm{d}\Omega$ the field equations derived from $\hat{\Omega}$ will have the same structure as those associated with Ω. However, the additional term $\mathrm{d}\Sigma$ in general will change the rank of Ω, and therefore the properties of the associated wave fronts. Thus, if we want $\hat{\Omega}$ to have the same rank as Ω we have to require $\mathrm{d}\Sigma = 0$ for a canonical transformation. In the case of Carathéodory's canonical theory we shall call a transformation (5,120) canonical, if there exists a function $\hat{H}_{\mathrm{c}} = \hat{H}_{\mathrm{c}}(\hat{x}, \hat{z}, \hat{p})$ such that

$$\hat{\Omega}_{\mathrm{c}} = -\hat{H}_{\mathrm{c}}^{-1}\,\hat{\theta}^1 \wedge \hat{\theta}^2 = -H_{\mathrm{c}}^{-1}\,\theta^1 \wedge \theta^2 = \Omega_{\mathrm{c}}\,, \qquad \hat{\theta}^{\mu} = -\hat{H}_{\mathrm{c}}\,\mathrm{d}\hat{x}^{\mu} + \hat{p}^{\mu}_{a}\,\mathrm{d}\hat{z}^a\,. \tag{5,121}$$

We here are not interested in the most general canonical transformation, but in the following special point transformation [E. Hölder, 1939]:

$$x^{\mu} \to \hat{x}^{\mu} = \hat{X}^{\mu}(x, z)\,, \qquad z^a \to \hat{z}^a = z^a\,. \tag{5,122}$$

Such transformations have the following properties and implications: If $z^a = f^a(x)$ is a smooth function, then we get

$$\mathrm{d}\hat{x}^{\mu}/\mathrm{d}x^{\nu} = \partial_{\nu}\hat{X}^{\mu} + \partial_a\hat{X}^{\mu}v^a_{\nu} =: A^{\mu}_{\nu}\,, \qquad v^a_{\mu} = \partial_{\mu}z^a(x)\,. \tag{5,123}$$

Since $\hat{z}^a(\hat{x}) = z^a(x)$, we have

$$v^a_{\mu} = \frac{\mathrm{d}z^a}{\mathrm{d}x^{\mu}} = \frac{\mathrm{d}\hat{z}^a}{\mathrm{d}\hat{x}^{\rho}}\frac{\mathrm{d}\hat{x}^{\rho}}{\mathrm{d}x^{\mu}} = \hat{v}^a_{\rho}A^{\rho}_{\mu} \tag{5,124}$$

and the eqs. (5,122–124) imply

$$\hat{v}^a_{\mu}\,\mathrm{d}\hat{x}^{\mu} = v^a_{\mu}\,\mathrm{d}x^{\mu} + \hat{v}^a_{\mu}\partial_b\hat{X}^{\mu}\omega^b\,, \tag{5,125}$$

because

$$\hat{v}^a_{\mu}\,\mathrm{d}\hat{x}^{\mu} = |A|^{-1}\,v^a_{\rho}\,\bar{A}^{\rho}_{\mu}\,(\partial_{\nu}\hat{X}^{\mu}\,\mathrm{d}x^{\nu} + \partial_b\hat{X}^{\mu}\,\mathrm{d}z^b)$$
$$= |A|^{-1}\,v^a_{\rho}\,\bar{A}^{\rho}_{\mu}\,(A^{\mu}_{\nu}\,\mathrm{d}x^{\nu} + \partial_b\hat{X}^{\mu}\,\omega^b)\,.$$

With $\mathrm{d}\hat{z}^a = \mathrm{d}z^a$ we therefore obtain for $\hat{\omega}^a = \mathrm{d}\hat{z}^a - \hat{v}^a_{\mu}\,\mathrm{d}\hat{x}^{\mu}$

$$\hat{\omega}^a = (\delta^a_b - \hat{v}^a_{\mu}\,\partial_b\hat{X}^{\mu})\omega^b\,, \tag{5,126}$$

which shows that the ideal $\mathrm{I}[\omega^a]$ is transformed into itself: If the ω^b vanish, then the $\hat{\omega}^a$ vanish, too.

Furthermore, we have

$$\begin{aligned}\hat{\theta}^{\mu} &= -\hat{H}_c\,\mathrm{d}\hat{x}^{\mu} + \hat{p}^{\mu}_a\,\mathrm{d}\hat{z}^a \\ &= -\hat{H}_c(\partial_\rho\hat{X}^{\mu}\,\mathrm{d}x^{\rho} + \partial_a\hat{X}^{\mu}\,\mathrm{d}z^a) + \hat{p}^{\mu}_a\,\mathrm{d}z^a \\ &= -\hat{H}_c\,\partial_\rho\hat{X}^{\mu}\,\mathrm{d}x^{\rho} + (-\hat{H}_c\,\partial_a\hat{X}^{\mu} + \hat{p}^{\mu}_a)\,\mathrm{d}z^a\,,\end{aligned}$$

and therefore

$$\begin{aligned}\hat{\Omega}_c &= -\hat{H}_c^{-1}\,\hat{\theta}^1 \wedge \hat{\theta}^2 = -\hat{H}_c\,|(\partial_\rho\hat{X}^{\mu})|\,\mathrm{d}x^1 \wedge \mathrm{d}x^2 + \partial_\rho\bar{\hat{X}}^1\,(-\hat{H}_c\,\partial_a\hat{X}^{\rho} + \hat{p}^{\rho}_a)\,\mathrm{d}z^a \wedge \mathrm{d}x^2 \\ &\quad + \partial_\rho\bar{\hat{X}}^2(-\hat{H}_c\,\partial_a\hat{X}^{\rho} + \hat{p}^{\rho}_a)\,\mathrm{d}x^1 \wedge \mathrm{d}z^a + \cdots \\ &= -H_c^{-1}\,\theta^1 \wedge \theta^2\,.\end{aligned}$$

Comparing coefficients of $\mathrm{d}x^1 \wedge \mathrm{d}x^2$ and $\mathrm{d}z^a \wedge \mathrm{d}x^{\mu}$ gives

$$H_c = \hat{H}_c|(\partial_\rho\hat{X}^{\mu})| \tag{5,127}$$

and

$$p^{\mu}_a = \partial_\rho\bar{\hat{X}}^{\mu}\,(-\hat{H}_c\,\partial_a\hat{X}^{\rho} + \hat{p}^{\rho}_a)\,, \tag{5,128}$$

or

$$|(\partial_\rho\hat{X}^{\nu})|\,\hat{p}^{\mu}_a = H_c\,\partial_a\hat{X}^{\mu} + \partial_\rho\hat{X}^{\mu}p^{\rho}_a\,.$$

The transformation properties of L and π^{μ}_a are obtained as follows:
With

$$\begin{aligned}\hat{a}^{\mu} &= \hat{L}\,\mathrm{d}\hat{x}^{\mu} + \hat{\pi}^{\mu}_a\hat{\omega}^a \\ &= \hat{L}[\partial_\rho\hat{X}^{\mu}\,\mathrm{d}x^{\rho} + \partial_a\hat{X}^{\mu}(\omega^a + v^a_\rho\,\mathrm{d}x^{\rho})] + \hat{\pi}^{\mu}_a(\delta^a_b - \hat{v}^a_\rho\,\partial_b\hat{X}^{\rho})\,\omega^b \\ &= \hat{L}\,A^{\mu}_\rho\,\mathrm{d}x^{\rho} + (\hat{\pi}^{\mu}_a - \hat{T}^{\mu}_\rho\,\partial_a\hat{X}^{\rho})\,\omega^a\,,\end{aligned}$$

$$\hat{T}^{\mu}_\rho = \hat{\pi}^{\mu}_a\hat{v}^a_\rho - \delta^{\mu}_\rho\hat{L}\,,$$

we get from $\hat{L}^{-1}\hat{a}^1 \wedge \hat{a}^2 = L^{-1}a^1 \wedge a^2$:

$$L = \hat{L}|A|\,, \qquad A = (A^{\mu}_\nu)\,, \tag{5,129}$$

and

$$\pi^{\mu}_a = \bar{A}^{\mu}_\rho(\hat{\pi}^{\rho}_a - \hat{T}^{\rho}_\nu\,\partial_a\hat{X}^{\nu})\,,$$

or

$$|A|\,(\hat{\pi}^{\mu}_{a} - \hat{T}^{\mu}_{\rho}\,\partial_a \hat{X}^{\rho}) = A^{\mu}_{\rho}\,\pi^{\rho}_{a}\,. \tag{5,130}$$

If we denote by

$$x^{\mu} = X^{\mu}(\hat{x}, \hat{z})\,, \qquad z^{a} = \hat{z}^{a} \tag{5,131}$$

the inverse of the transformation (5,122), then we get the corresponding inverse transformations for the other quantities by interchanging in the above formulae all quantities with a "hat" ^ by the corresponding ones without a hat and vice versa. For instance, the inverses of the formulae (5,127) and (5,128) are

$$\hat{H}_c = H_c|(\hat{\partial}_{\rho} X^{\mu})|\,, \qquad \hat{\partial}_{\rho} := \partial/\partial \hat{x}^{\rho}\,, \tag{5,132}$$

$$|(\hat{\partial}_{\rho} X^{\nu})|\,p^{\mu}_{a} = \hat{H}_c \partial_a X^{\mu} + \hat{\partial}_{\rho} X^{\mu} \hat{p}^{\rho}_{a}\,. \tag{5,133}$$

Because of the identities

$$x^{\mu} = X^{\mu}[\hat{X}(x, z), z]\,, \qquad \hat{x}^{\mu} = \hat{X}^{\mu}[X(x, z), z]\,,$$

we have the relations

$$\hat{\partial}_{\rho} X^{\mu}\,\partial_{\nu} \hat{X}^{\rho} = \delta^{\mu}_{\nu}\,, \qquad \hat{\partial}_{\rho} X^{\mu}\,\partial_a \hat{X}^{\rho} + \partial_a X^{\mu} = 0\,, \qquad \partial_{\rho} \hat{X}^{\mu}\,\partial_a X^{\rho} + \partial_a \hat{X}^{\mu} = 0\,. \tag{5,134}$$

The importance of the transformations (5,122) lies in the following remarkable special case [E. Hölder, 1939]:

$$\hat{x}^{1} = x^{1}\,, \qquad \hat{x}^{2} = \sigma^{2} = S^{2}(x, z)\,, \qquad \hat{z}^{a} = z^{a}\,. \tag{5,135}$$

Since S^2 obeys eqs. (5,69b), $p^{\rho}_{a}\,\partial_{\rho} S^2 + H_c\,\partial_a S^2 = 0$, a comparison with eqs. (5,128) gives

$$\partial_2 S^2 \hat{p}^{1}_{a} = p^{1}_{a}\,, \qquad \hat{p}^{2}_{a} = 0\,. \tag{5.136}$$

The equations $\hat{p}^{2}_{a} = 0$ have a number of very interesting consequences: For $\hat{S}^{2}_{\mu} = \hat{p}^{2}_{a}\hat{v}^{a}_{\mu} - \delta^{2}_{\mu}\hat{H}_c$ we get $\hat{S}^{2}_{\mu} = -\delta^{2}_{\mu}\hat{H}_c$ and therefore it follows from $\hat{L}\hat{p}^{2}_{a} = \hat{S}^{2}_{\mu}\hat{\pi}^{\mu}_{a}$ – compare eq. (5,13) – that

$$\hat{\pi}^{2}_{a} = 0\,. \tag{5,137}$$

This implies $\hat{T}^{2}_{\mu} = -\delta^{2}_{\mu}\hat{L}$ and we have

$$\hat{H}_c = -\hat{L}^{-1}|(\hat{T}^{\mu}_{\nu})| = \hat{T}^{1}_{1} = \hat{\pi}^{1}_{a}\hat{v}^{a}_{1} - \hat{L} = \hat{K}\,. \tag{5,138}$$

Thus, in the E. Hölder frame the Hamilton function $\hat{H}_c$ is equal to the canonical energy density $\hat{T}^{1}_{1}$, if x^1 is the time coordinate, which it is in most physical systems!

Combined with $\hat{H}_c\hat{\pi}^1_a = \hat{T}^1_\mu \hat{p}^\mu_a = \hat{T}^1_1 \hat{p}^1_a$ eq. (5,138) gives

$$\hat{\pi}^1_a = \hat{p}^1_a \,. \tag{5,139}$$

Furthermore, the complicated canonical eqs. (5,26c) and (5,67) become very simple and familiar: Because $\hat{K} = \hat{T}^1_1 = \hat{H}_c$ it follows from

$$\hat{K}^{-1}\hat{v}^a_\mu \hat{T}^\mu_1 = \hat{K}^{-1}\hat{v}^a_1 \hat{T}^1_1 = \partial \hat{H}_c/\partial \hat{p}^1_a$$

that

$$\mathrm{d}\hat{z}^a/\mathrm{d}\hat{x}^1 = \partial \hat{H}_c/\partial \hat{p}^1_a \,. \tag{5,140}$$

Because $\hat{p}^2_a = 0$, $\hat{T}^2_1 = 0$, we have $\partial \hat{p}^2_a/\partial \hat{x}^\mu = 0$ and the coefficients $C^{b\mu}_{a1}$ and C^μ_a in the eqs. (5,67) take the form

$$\hat{C}^{b1}_{a1} = \hat{H}_c \delta^b_a \,, \qquad \hat{C}^{b2}_{a1} = 0 \,, \qquad \hat{C}^\mu_a = 0 \,. \tag{5,141}$$

We therefore get as the second set of Carathéodory's canonical field equations in the E. Hölder frame the "mechanical" equations

$$\mathrm{d}\hat{p}^1_a/\mathrm{d}\hat{x}^1 = -\partial \hat{H}_c/\partial \hat{z}^a \,. \tag{5,142}$$

Before going into the interpretation of the above formulae and illustrating them be examples, let us derive the transformation properties of the other quantities:
The matrix $A = (A^\mu_\rho)$, eq. (5,123), has the form

$$\begin{pmatrix} 1 & 0 \\ \Delta^2_1 & \Delta^2_2 \end{pmatrix}, \tag{5,143}$$

so that

$$\hat{A} = A^{-1} = \frac{1}{\Delta^2_2}\begin{pmatrix} \Delta^2_2 & 0 \\ -\Delta^2_1 & 1 \end{pmatrix} \tag{5,144}$$

and it follows from eqs. (5,124) that

$$\hat{v}^a_1 = v^a_1 - (\Delta^2_1/\Delta^2_2)\, v^a_2 \,, \qquad \hat{v}^a_2 = v^a_2/\Delta^2_2 \,. \tag{5,145}$$

Since $\Delta^2_\mu = \partial_\mu S^2 + v^a_\mu \partial_a S^2$ we obtain from eqs. (5,145)

$$\begin{aligned} \Delta^2_1 &= (\partial_1 S^2 + \hat{v}^a_1\, \partial_a S^2)\,(1 - \hat{v}^a_2\, \partial_a S^2)^{-1} \,, \\ \Delta^2_2 &= \partial_2 S^2 (1 - \hat{v}^a_2\, \partial_a S^2)^{-1} \,. \end{aligned} \tag{5,146}$$

The inverse of the eqs. (5,130) is

$$\hat{A}^{\mu}_{\rho}\,\hat{\pi}^{\rho}_{a} = |\hat{A}|\,(\pi^{\mu}_{a} - T^{\mu}_{\rho}\,\partial_a X^{\rho})\,,$$

which in our special case implies

$$\begin{aligned}\hat{\pi}^{1}_{a} &= (\pi^{1}_{a} + T^{1}_{2}\,\partial_a S^2/\partial_2 S^2)/\Delta^{2}_{2} \\ &= \pi^{1}_{b}(\delta^{b}_{a} + v^{b}_{2}\,\partial_a S^2/\partial_2 S^2)/\Delta^{2}_{2}\,,\end{aligned} \tag{5,147}$$

because, according to eqs. (5,134), $\partial_a X^2 = -\partial_a S^2/\partial_2 S^2$. The transformation (5,130) itself takes the form

$$\begin{aligned}\pi^{1}_{a} &= \Delta^{2}_{2}(\hat{\pi}^{1}_{a} - \hat{T}^{1}_{2}\,\partial_a S^2) = \Delta^{2}_{2}\,\hat{\pi}^{1}_{b}(\delta^{b}_{a} - \hat{v}^{b}_{2}\,\partial_a S^2)\,,\\ \pi^{2}_{a} &= -\Delta^{2}_{1}\,\hat{\pi}^{1}_{b}(\delta^{b}_{a} - \hat{v}^{b}_{2}\,\partial_a S^2) + \hat{L}\,\partial_a S^2\,,\end{aligned} \tag{5,148}$$

from which we get

$$\Delta^{2}_{\mu}\pi^{\mu}_{a} = \Delta^{2}_{2}\hat{L}\,\partial_a S^2 = L\,\partial_a S^2\,. \tag{5,149}$$

The eqs. (5,148–149) can be used in order to prove the consistency relation

$$\begin{aligned}\hat{\pi}^{2}_{a} &= \frac{\partial\hat{L}}{\partial\hat{v}^{a}_{2}} = \frac{\partial}{\partial\hat{v}^{a}_{2}}(L/\Delta^{2}_{2}) \\ &= \left(\pi^{\nu}_{b}\frac{\partial v^{b}_{\nu}}{\partial\hat{v}^{a}_{2}}\right)\Big/\Delta^{2}_{2} - L\frac{\partial\Delta^{2}_{2}}{\partial\hat{v}^{a}_{2}}\Big/(\Delta^{2}_{2})^2 = 0\,.\end{aligned}$$

Because $\hat{\theta}^2 = -\hat{H}_c\,d\sigma^2$, our canonical form $\Omega_c = \hat{\Omega}_c$ becomes in the E. Hölder frame

$$\Omega_c = \hat{\theta}^1 \wedge d\sigma^2\,,\qquad \hat{\theta}^1 = -\hat{H}_c\,d\hat{x}^1 + \hat{p}^{1}_{a}\,d\hat{z}^{a}\,, \tag{5,150}$$

which shows that on the "characteristic" hypersurfaces $S^2(x, z) = \sigma^2 = $ const. the system behaves like a mechanical one, in accordance with all the formulae derived above and which are valid on these hypersurfaces.

If we take $\hat{S}^2 = \sigma^2 = \hat{x}^2$, the CHJ equation in the E. Hölder frame follows from

$$d\hat{S}^1 \wedge d\sigma^2 = \hat{\theta}^1 \wedge d\sigma^2\,, \tag{5,151}$$

so that

$$\hat{\partial}_1\hat{S}^1 + \hat{H}_c = 0\,,\qquad \hat{p}^{1}_{a} = \partial_a\hat{S}^1\,, \tag{5,152}$$

again as in mechanics!

From

$$d\hat{S}^1 \wedge d\sigma^2 = d\hat{S}^1 \wedge d\hat{S}^2 = dS^1 \wedge dS^2$$

we obtain the transformation properties

$$\hat{\partial}_1 \hat{S}^1 = |(\partial_\mu S^\nu)|/\partial_2 S^2 ,$$
$$\partial_a \hat{S}^1 = (\partial_2 S^2 \, \partial_a S^1 - \partial_2 S^1 \, \partial_a S^2)/\partial_2 S^2 . \tag{5,153}$$

Furthermore, if S^μ and $\hat{S}^1$ depend on a parameter a, we have, instead of eq. (5,151),

$$\mathrm{d}\hat{S}^1 \wedge \mathrm{d}\sigma^2 = \hat{\theta}^1 \wedge \mathrm{d}\sigma^2 + \frac{\partial \hat{S}^1}{\partial a} \mathrm{d}a \wedge \mathrm{d}\sigma^2$$
$$= -H_c^{-1} \, \theta^1 \wedge \theta^2 + G^\mu \, \mathrm{d}a \wedge \mathrm{d}\Sigma_\mu , \tag{5,154}$$
$$G^\mu = \frac{\partial S^\rho}{\partial a} \bar{\Delta}^\mu_\rho .$$

The first equality in these equations implies

$$\hat{G}^1 = \partial \hat{S}^1/\partial a , \qquad \hat{G}^2 = 0 , \tag{5,155}$$

where $\hat{G}^1 = \text{const.}$ on the extremals in the E. Hölder frame (compare section 5.6). Because $\hat{\theta}^1 \wedge \mathrm{d}\sigma^2 = -H_c^{-1} \, \theta^1 \wedge \theta^2$ we further conclude from eq. (5,154) that

$$\frac{\partial \hat{S}^1}{\partial a} \mathrm{d}a \wedge \mathrm{d}S^2 = G^\mu \, \mathrm{d}a \wedge \mathrm{d}\Sigma_\mu ,$$

from which we get

$$G^1 = \Delta^2_2 \, \partial \hat{S}^1/\partial a , \qquad G^2 = -\Delta^2_1 \, \partial \hat{S}^1/\partial a . \tag{5,156}$$

An application of the above formulae can proceed as follows:
Since

$$(\partial_\rho \hat{X}^\mu) = \begin{pmatrix} 1 & \partial_1 S^2 \\ 0 & \partial_2 S^2 \end{pmatrix} ,$$

we have

$$(\hat{\partial}_\rho X^\mu) = (\partial_\rho \hat{X}^\mu)^{-1} = \frac{1}{\partial_2 S^2} \begin{pmatrix} \partial_2 S^2 & -\partial_1 S^2 \\ 0 & 1 \end{pmatrix} ,$$

and since $\partial_a X^2 = -\partial_a S^2/\partial_2 S^2$, we obtain from eqs. (5,133)

$$p^1_a = \partial_2 S^2 p_a , \qquad p^2_a = -\hat{H}_c \, \partial_a S^2 - \partial_1 S^2 p_a , \qquad p_a := \hat{p}^1_a . \tag{5,157}$$

Inserting these values for p^1_a and p^2_a into eq. (5,127),

$$\partial_2 S^2 \hat{H}_c = H_c(x, z = \hat{z}, p_a^1 = \partial_2 S^2 p_a, p_a^2 = -\hat{H}_c\, \partial_a S^2 - \partial_1 S^2 p_a)\,, \tag{5,158}$$

$$x^1 = \hat{x}^1 := t\,, \qquad x^2 = X^2(t, \sigma^2, \hat{z})\,,$$

we obtain an algebraic equation for $\hat{H}_c(t, z, p)$, the Hamilton function for the effective "mechanical" problem in the surfaces $S^2 = \text{const}$. The integration of the canonical eqs. (5,140) and (5,142) may be simplified by "energy" conservation in the surfaces $S^2(x, z) = \text{const.}$:
Suppose the function $\hat{H}_c$ does not depend explicitly on the coordinate t, then we have $\hat{H}_c = \hat{E} = \text{const.}$ in the surfaces $S^2(x, z) = \text{const.}$
Examples:
(i) $H_c = \frac{1}{2}((p^1)^2 - (p^2)^2) + V(z)$. Here we have

$$p^1 = \partial_2 S^2 p\,, \qquad p^2 = -\hat{H}_c\, \partial_z S^2 - \partial_1 S^2 p$$

and eq. (5,158) becomes

$$\partial_2 S^2 \hat{H}_c = \tfrac{1}{2}(\partial_2 S^2 p)^2 - \tfrac{1}{2}(\hat{H}_c \partial_z S^2 + \partial_1 S^2 p)^2 + V(z)\,. \tag{5,159}$$

This is a quadratic equation for $\hat{H}_c$. However, in order to calculate the derivatives $\partial \hat{H}_c/\partial p$ and $\partial \hat{H}_c/\partial z$ we can differentiate eq. (5,159) directly and obtain

$$\frac{\partial \hat{H}_c}{\partial p}[\partial_2 S^2 + (\hat{H}_c\, \partial_z S^2 + p\, \partial_1 S^2)\, \partial_z S^2] = p[(\partial_2 S^2)^2 - (\partial_1 S^2)^2] - \hat{H}_c\, \partial_1 S^2\, \partial_z S^2\,, \tag{5,160}$$

$$\frac{\partial \hat{H}_c}{\partial z}[\partial_2 S^2 + (\hat{H}_c\, \partial_z S^2 + p\, \partial_1 S^2)\, \partial_z S^2]$$

$$= \partial_z V(z) + \partial_2 \partial_z S^2 (p^2 \partial_2 S^2 - \hat{H}_c) - (\hat{H}_c\, \partial_z S^2 + \partial_1 S^2 p)(\hat{H}_c\, \partial_z^2 S^2 + p\, \partial_1 \partial_z S^2)\,. \tag{5,161}$$

As to the integration of the equations

$$\dot{z} = \partial \hat{H}_c/\partial p\,, \qquad \dot{p} = -\partial \hat{H}_c/\partial z\,,$$

the following method seems to be of interest, provided $\partial_1 S^2 = 0$, $\partial_2 \partial_z S^2 = 0$:
From

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{\mathrm{d}p}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\mathrm{d}p}{\mathrm{d}z}\frac{\partial \hat{H}_c}{\partial p} = -\frac{\partial \hat{H}_c}{\partial z}$$

and the eqs. (5,160, 161) we obtain

$$\frac{\mathrm{d}p}{\mathrm{d}z} p = [-\partial_z V + \tfrac{1}{2}\hat{H}_c^2\, \partial_z (\partial_z S^2)^2]\, (\partial_2 S^2)^{-2}\,.$$

If the r.h. side is a function of z only, for instance, because $\hat{H}_c = \hat{E} = \text{const.}$, we can calculate $p(z)$ by integration

$$\tfrac{1}{2}p^2(z) = \int^z \mathrm{d}\bar{z}\,[-\partial_z V + \tfrac{1}{2}\hat{H}_c^2\,\partial_z(\partial_z S^2)^2]\,(\partial_2 S^2)^{-2}\,. \tag{5,162}$$

If the factor $(\partial_2 S^2)^2\,(\partial_2 S^2 + H_c(\partial_z S^2)^2)^{-1}$ in eq. (5,160) does only depend on z, too, and not on t, we can integrate

$$\mathrm{d}t/\mathrm{d}z = p^{-1}(z)\,(\partial_2 S^2)^{-2}\,(\partial_2 S^2 + \hat{H}_c\,(\partial_z S^2)^2)\,. \tag{5,163}$$

This method works for the special solution (5,82) for which

$$\partial_2 S^2 = \frac{1}{\sqrt{2}}\,\omega^{1/2}A\,, \qquad \hat{H}_c = \frac{1}{\sqrt{2}}\,\omega^{-1/2}A\,, \qquad \partial_z S^2 = -\sqrt{2}A^{-1}k\omega^{1/2}(A^2 - 2V)^{1/2}\,,$$

so that

$$\tfrac{1}{2}\hat{H}_c^2\partial_z(\partial_z S^2)^2 = -k^2\partial_z V$$

and eq. (5,162) becomes

$$\tfrac{1}{2}p^2(z) = \omega A^{-2}(A^2 - 2V(z))\,,$$

where the constant of integration has been chosen appropriately.

Since

$$\partial_2 S^2 + \hat{H}_c(\partial_z S^2)^2 = \frac{1}{\sqrt{2}}A\omega^{1/2}[1 + 2A^{-2}k^2(A^2 - 2V)]\,,$$

eq. (5,163) becomes

$$\mathrm{d}t/\mathrm{d}z = \omega^{-1}[1 + 2A^{-2}k^2(A^2 - 2V)](A^2 - 2V(z))^{-1/2}\,,$$

which gives

$$\omega t + \text{const.} = \int^z \mathrm{d}\bar{z}\,(A^2 - 2V(\bar{z}))^{-1/2} + 2k^2A^{-2}\int^z \mathrm{d}\bar{z}\,(A^2 - 2V(\bar{z}))^{1/2}\,. \tag{5,164}$$

Since $S^2(x, z) = \sigma^2 = \text{const.}$, we obtain from eqs. (5,82) and (5,84)

$$2k^2A^{-2}\int^z \mathrm{d}\bar{z}\,(A^2 - 2V(\bar{z}))^{1/2} = kx^2 - \sqrt{2}A^{-1}\omega^{-1/2}\sigma^2\,,$$

which, combined with the expression (5,164), gives

$$\int^{z} \mathrm{d}\bar{z}\,(A^2 - 2V(\bar{z}))^{-1/2} = \omega x^1 - kx^2 + \alpha\,, \qquad \alpha = \text{const.} \tag{5.165}$$

This equation, when solved for $z = f(\omega x^1 - kx^2)$, provides us with a solution of the original problem.

(ii) Relativistic string.

From eq. (5,85) we obtain

$$\partial_{(1)}S^2 = 0\,, \qquad \partial_{(2)}S^2 = \tfrac{1}{2}\,, \qquad \partial_\alpha S^2 = -(1-\omega^2\rho^2)^{1/2}\,\partial_\alpha\rho\,,$$
$$(\partial_\alpha\rho) = (0, \boldsymbol{e}_\rho, 0)\,, \qquad \boldsymbol{e}_\rho = (x^1, x^2)/\rho\,, \tag{5,166}$$

and therefore the eqs. (5,157) take here the form

$$p^1_\alpha = \tfrac{1}{2}p_\alpha\,, \qquad p^2_\alpha = -\hat{H}_c\,\partial_\alpha S^2\,. \tag{5,167}$$

From eqs. (5,166, 167) we get

$$(\partial S^2)^2 = -(1-\omega^2\rho^2)\,, \qquad p\cdot(\partial S) = \boldsymbol{p}\cdot\boldsymbol{e}_\rho(1-\omega^2\rho^2)^{1/2}\,,$$
$$(\partial S^2) := (\partial_0 S^2, \ldots, \partial_3 S^2)\,,$$

and by using the Hamilton function (5,59) the third power of eq. (5,158) becomes

$$(\partial_{(2)}S^2)^3\,\hat{H}^3_c = H^3_c = -\tfrac{1}{8}(\partial_{(2)}S^2)^2\,\hat{H}^2_c\,[(p\cdot p)\,(\partial S^2)^2 - (p\cdot(\partial S^2))^2]\,,$$

or

$$\hat{H}_c = \tfrac{1}{4}[(p\cdot p) + (\boldsymbol{p}\cdot\boldsymbol{e}_\rho)^2]\,(1-\omega^2\rho^2)\,. \tag{5,168}$$

In the following it is convenient to use polar coordinates ρ and θ in the (x^1, x^2)-plane. Since $p_\rho = (\boldsymbol{p}\boldsymbol{e}_\rho)$, $(p^1)^2 + (p^2)^2 = p^2_\rho + p^2_\theta/\rho^2$, the Hamilton function (5,168) takes the form

$$\hat{H}_c = \tfrac{1}{4}[(p_0)^2 - (p_3)^2 - (p_\theta)^2/\rho^2](1-\omega^2\rho^2)\,, \tag{5,169}$$

i.e., the p_ρ-term has dropped out of $\hat{H}_c$. The "mechanical" equations of motion are

$$\dot{x}^0 = \tfrac{1}{2}p_0(1-\omega^2\rho^2)\,, \qquad \dot{x}^3 = -\tfrac{1}{2}p_3(1-\omega^2\rho^2)\,,$$
$$\dot{\rho} = 0\,, \qquad \dot{\theta} = -\tfrac{1}{2}p_\theta(1-\omega^2\rho^2)/\rho^2\,, \tag{5,170}$$

and

$$\dot{p}_0 = 0 \,, \qquad \dot{p}_3 = 0 \,, \qquad \dot{p}_\theta = 0 \,,$$
$$\dot{p}_\rho = -\tfrac{1}{2} p_\theta^2 \rho^{-3} (1-\omega^2\rho^2) + 2\omega^2 \rho \hat{H}_c (1-\omega^2\rho^2)^{-1}$$
$$= 2\rho^2(\omega^2 \hat{H} - \dot{\theta}^2)(1-\omega^2\rho^2)^{-1} \,. \tag{5,171}$$

We fix the constant p_0 such that $\dot{x}^0 = 1$, i.e. $x^0 = \tau^1$ and $p_0 = 2(1-\omega^2\rho^2)^{-1}$. We further take $p_3 = 0$ and $x^3 = 0$. Since p_θ = const. and ρ = const., it follows from the last of the eqs. (5,170) that $\dot{\theta}$ is constant, too. We take $\dot{\theta} = \omega$, which implies $\hat{H}_c = 1$ and therefore, according to the last of eqs. (5,171), $\dot{p}_\rho = 0$, which is compatible with $p_\rho = 0$. We thus have the following special solution of the equations of motion (5,170, 171):

$$x^0(\tau^1) = \tau^1 \,, \qquad x^1(\tau^1) = \rho \cos \omega\tau^1 \,, \qquad x^2(\tau^1) = \rho \sin \omega\tau^1 \,, \qquad x^3 = 0 \,.$$

The τ^2 dependence of ρ we can derive with the help of eq. (5,62):

$$\partial_{(2)} x^\alpha = \tfrac{1}{4} H_c^{-2} g^{\alpha\beta} \left[p^1_\beta (p^1 \cdot p^2) - p^2_\beta (p^1 \cdot p^1) \right] .$$

According to eqs. (5,166, 167) we have

$$p^1 \cdot (\partial S^2) = 0 \,, \qquad p^1 \cdot p^1 = (1-\omega^2\rho^2)^{-1} \,.$$

Taking further into account that $H_c^2 = \frac{1}{4}$ and $\partial_\alpha S^2 = -(1-\omega^2\rho^2)^{1/2} \partial_\alpha \rho$ we get

$$(\partial_{(2)} x^\alpha) = (1-\omega^2\rho^2)^{1/2} \, (0, e_\rho, 0) \,, \tag{5,172}$$

which coincides with the τ^2-dependence of ρ in eq. (5,86)! The Lagrangian $\hat{L} = p_\alpha \hat{v}_1^\alpha - \hat{H}_c$ corresponding to the Hamilton function (5,169) is

$$\hat{L} = \hat{H}_c = [(\dot{x}^0)^2 - (\dot{x}^3)^2 - \rho^2 \dot{\theta}^2] \, (1-\omega^2\rho^2)^{-1} \,. \tag{5,173}$$

5.8. Embedding a given extremal into a system of CHJ wave fronts

I finally want to mention, how – at least in principle – a given extremal $z^a = f^a(x)$ can be weakly embedded in a geodesic field $S^\mu(x, z)$ of Carathéodory. The basic idea and the first proof is due to Boerner [1936]. It was essentially improved by Hölder [1939], elegantly put into the framework of differential forms by Lepage [1942b] and further discussed by Van Hove [1945b]:
Since the CHJ equation is one partial differential equation for two functions $S^1(x, z)$ and $S^2(x, z)$, one first tries to construct a function $S^2(x, z)$ which obeys the transversality condition

$$\mathring{H}_c \, \partial_\alpha S^2 + \mathring{p}^\rho_a \, \partial_\rho S^2 = 0 \tag{5,174}$$

in the point $(\mathring{x}, f(\mathring{x}))$ of the extremal, where $\mathring{p}^\rho_a = p^\rho_a(\mathring{x})$, $\mathring{H}_c = H_c(\mathring{x}, f(\mathring{x}), p(\mathring{x}))$. Having found such a $S^2(x, z)$, we can use it for a E. Hölder transformation as discussed above in order to reduce the problem to a "mechanical" one which then can be treated in the standard way mentioned in chapter 2. The (local!) construction of $S^2(x, z)$ can – in principle – proceed as follows: Consider the plane

$$x^\mu = \mathring{x}^\mu + (\mathring{p}^\mu_a/\mathring{H}_c)(z^a - f^a(\mathring{x})) \tag{5,175}$$

with "running" variables (x, z), through the point $(\mathring{x}, f(\mathring{x}))$ of the extremal. Since

$$\left.\frac{dx^\mu}{d\mathring{x}^\nu}\right|_{z=f(\mathring{x})} = \delta^\mu_\nu - (\mathring{p}^\mu_a/\mathring{H}_c)\,\partial_\nu f^a(\mathring{x}) = -\mathring{S}^\mu_\nu/\mathring{H}_c\,,$$

and since $|\mathring{S}| = -H_cL$, compare eq. (5,11), the eq. (5,175) can be solved for $\mathring{x}$ in a neighborhood of $(\mathring{x}, f(\mathring{x}))$ if $\mathring{H}_c\mathring{L} \neq 0$. Suppose this is the case and let $\mathring{x}^\mu = \xi^\mu(x, z)$ be the result. If $s^2(\xi)$ is a smooth function of the variables ξ^μ, $\mu = 1, 2$, then

$$S^2(x, z) = s^2[\xi(x, z)] \tag{5,176}$$

fulfills the transversality conditions (5,174). This can be seen as follows:
From eq. (5,176) we get

$$\partial_\nu S^2 = (\partial s^2/\partial\xi^\rho)\,\partial_\nu\xi^\rho\,, \qquad \partial_a S^2 = (\partial s^2/\partial\xi^\rho)\,\partial_a\xi^\rho\,. \tag{5,177}$$

The derivatives $\partial_\nu\xi^\rho$, $\partial_a\xi^\rho$ at $z^a = f^a(\mathring{x})$ can be calculated from the identity

$$\xi^\rho = x^\rho - [p^\rho_b(\xi)/H(\xi, z(\xi), p(\xi))]\,(z^b - f^b(\xi))\,, \qquad \xi = \xi(x, z)\,,$$

from which we get

$$\partial_\mu\xi^\rho = \delta^\rho_\mu + (\mathring{p}^\rho_b/\mathring{H}_c)\,\mathring{v}^b_\nu\,\partial_\mu\xi^\nu\,,$$

$$\partial_a\xi^\rho = -\mathring{p}^\rho_a/\mathring{H}_c + (\mathring{p}^\rho_b/\mathring{H}_c)\,\mathring{v}^b_\nu\,\partial_a\xi^\nu\,.$$

These equations can be rewritten as

$$\partial_\mu\xi^\rho\mathring{S}^\nu_\rho = -\delta^\nu_\mu\mathring{H}_c\,, \qquad \partial_a\xi^\rho\mathring{S}^\nu_\rho = \mathring{p}^\nu_a\,,$$

or, since $T^\lambda_\nu S^\nu_\rho = -\delta^\lambda_\rho\,|T|$,

$$\partial_\mu\xi^\rho|\mathring{T}| = \mathring{H}_c\mathring{T}^\rho_\mu\,, \qquad \text{or} \qquad \partial_\mu\xi^\rho = -\mathring{T}^\rho_\mu/\mathring{L}\,, \tag{5,178a}$$

$$\partial_a\xi^\rho|\mathring{T}| = -\mathring{T}^\rho_\nu\mathring{p}^\nu_a\,, \qquad \text{or} \qquad \partial_a\xi^\rho = \mathring{\pi}^\rho_a/\mathring{L}\,, \tag{5,178b}$$

where the relation (5,8) has been used. This same relation implies that the expressions (5,178) for $\partial_\mu\xi^\rho$ and $\partial_a\xi^\rho$, when combined with the eqs. (5,177) fulfill the transversality conditions (5,174).

The simplest choice for s^2 is $s^2(\xi) = \xi^2$ which gives $\partial_\mu s^2 = \partial_\mu\xi^2$, $\partial_a s^2 = \partial_a\xi^2$. The expressions (5,178) suggest the following linear ansatz for $S^2(x, z)$ in the neighborhood of an extremal $z^a = f^a(x)$:

$$S^2(x,z) = x^2 + (\mathring{\pi}_a^2(x)/\mathring{L})(z^a - f^a(x)), \tag{5,179}$$

with the derivatives

$$\partial_\mu S^2 = \partial_\mu(\mathring{\pi}_a^2/\mathring{L})(z^a - f^a(x)) - \mathring{T}_\mu^2/\mathring{L}, \qquad \partial_a S^2 = \mathring{\pi}_a^2/\mathring{L},$$

which for $z^a = f^a(x)$ are the same as those obtained above, eq. (5,178).

5.9. Bibliographical notes

Carathéodory's important paper on a canonical theory for fields appeared in 1929. As far as I know it has not received any attention in the physics community. It was first taken up by Boerner [1936] who summarized it and proved that any extremal could be embedded in a system of CHJ wave fronts. The most important analysis of Carathéodory's theory is due to E. Hölder [1939] who recognized its close relationship to Huygens' principle and who discovered the canonical transformation which reduces the dynamical equations to corresponding mechanical ones.

The special role of Carathéodory's theory within a more general canonical framework became clear through the work of Lepage [1936a, b, 1941, 1942a, b] who made clear, in terms of differential forms, that Carathéodory's theory is unique in the sense that it is the only theory which has a fundamental canonical form with minimal rank $r = m$, where m is the number of independent variables. This property implies that the associated transversal wave fronts have the same dimension n as the number of dependent variables.

Lepage's results were discussed by Boerner [1940b, 1953]. See also the textbooks by Funk [1970] and Rund [1973].

An evaluation of the theory in the light of more modern differential geometry (global aspects, algebraic topology, foliations etc.) is due to Dedecker [1953, 1957, 1977, 1978, 1980 (with Tulczyjew)]. His papers and results, which are not easy to understand, at least by a physicist, are waiting for an interpretation in physical terms. The same expectation applies to the related work by Vinogradov [1977, 1978] and Kupershmidt* [1980].

There is no doubt that the wave fronts, transversal to the extremals, play a very central role in Carathéodory's approach to the calculus of variations [1930, 1935, 1937], in the tradition of the ideas of Huygens, Cauchy, Hamilton, Jacobi and Weierstraß [E. Hölder et al., 1966]. A similar emphasis on the "wave" aspects of dynamical systems can be found in the work of Vessiot [1906, 1909, 1912, 1913, 1914], reviewed by Juvet [1937].

In view of the extremely strong influence the HJ theory of mechanics has had on the development of quantum mechanics, perhaps HJ theories for fields will have some influence on the future developments of quantum field theories, too!

6. Canonical theories for fields with m independent variables

6.1. The general case

The extension of the discussion in the previous chapters for systems with two independent variables to systems with $m > 2$ independent ones is straightforward and does not bring anything new concep-

*See, however, the references added in proof!

tually. I shall, therefore, be rather brief here and shall only indicate some essential points: If

$$\omega = L\,\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^m\,, \qquad L = L(x, z, v)\,, \tag{6,1}$$

is the Lagrangian m-form which defines the system, then it belongs to an equivalence class of m-forms, the most general representative Ω of which consists of ω plus a linear combination of all m-forms – with arbitrary coefficients – which are obtained from $\mathrm{d}x^1 \wedge \cdots \wedge \mathrm{d}x^m$ by replacing $1, 2, \ldots$, or all of the $\mathrm{d}x^\mu$ by the 1-forms $\omega^a = \mathrm{d}z^a - v^a_\mu\,\mathrm{d}x^\mu$ which vanish on the extremals where $v^a_\mu = \partial_\mu f^a(x)$ and which generate an ideal $\mathrm{I}[\omega^a]$. To be more specific, let us assume that we have a system in Minkowski space M^4 with coordinates $x^0, \ldots, x^3$ (in our units the velocity of light c has the value 1). Then Ω has the form

$$\Omega = L\,\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3 + h^\mu_a\,\omega^a \wedge \mathrm{d}^3\Sigma_\mu + \tfrac{1}{4}h^{\mu\nu}_{ab}\,\omega^a \wedge \omega^b \wedge \mathrm{d}^2 S_{\mu\nu}$$

$$+ \frac{1}{3!}\,h_{abc;\mu}\,\omega^a \wedge \omega^b \wedge \omega^c \wedge \mathrm{d}x^\mu + \frac{1}{4!}\,h_{abcs}\,\omega^a \wedge \omega^b \wedge \omega^c \wedge \omega^s\,, \tag{6,2}$$

where

$$\mathrm{d}^3\Sigma_\mu = \frac{1}{3!}\,\varepsilon_{\mu\alpha\beta\gamma}\,\mathrm{d}x^\alpha \wedge \mathrm{d}x^\beta \wedge \mathrm{d}x^\gamma\,, \qquad \mathrm{d}^2 S_{\mu\nu} = \tfrac{1}{2}\varepsilon_{\mu\nu\alpha\beta}\,\mathrm{d}x^\alpha \wedge \mathrm{d}x^\beta\,.$$

Here $\varepsilon_{\mu\nu\rho\sigma}$ is the totally antisymmetric tensor with $\varepsilon_{0123} = +1$. The coefficients $h^{\mu\nu}_{ab}$ are antisymmetric in the indices (μ, ν) and (a, b) separately, the coefficients $h_{abc;\mu}$ completely antisymmetric in (a, b, c) etc. Thus, the term h_{abcs} can only occur if $n \geq 4$. The coefficients h^μ_a, $h^{\mu\nu}_{ab}$ etc. can be arbitrary functions of x, z and v!

As before, the coefficients h^μ_a are determined to be equal to $\pi^\mu_a = \partial L/\partial v^a_\mu$ by the requirement that $\mathrm{d}\Omega \equiv 0 \pmod{\mathrm{I}[\omega^a]}$!

The Legendre transformation $v^a_\mu \to p^\mu_a$, $L \to H$, is again implemented by inserting on the r.h. side of eq. (6,2) for ω^a the expression $\mathrm{d}z^a - v^a_\mu\,\mathrm{d}x^\mu$ and identifying H with the resulting negative coefficient of $\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3$, and the canonical momenta p^μ_a with the coefficients of $\mathrm{d}z^a \wedge \mathrm{d}^3\Sigma_\mu$ etc. If $h_{abc;\mu} = 0$, $h_{abcs} = 0$, we obtain

$$p^\mu_a = \pi^\mu_a - h^{\mu\nu}_{ab} v^b_\nu\,, \tag{6.3a}$$

$$H = \pi^\mu_a v^a_\mu - \tfrac{1}{2} h^{\mu\nu}_{ab} v^a_\mu v^b_\nu - L\,. \tag{6,3b}$$

In performing the calculations, the following identities are useful [see e.g. Misner et al., 1973, Box 5.3]:

$$\varepsilon^{\alpha\beta\gamma\mu}\,\mathrm{d}^3\Sigma_\mu = \mathrm{d}x^\alpha \wedge \mathrm{d}x^\beta \wedge \mathrm{d}x^\gamma\,, \qquad \mathrm{d}x^\rho \wedge \mathrm{d}^3\Sigma_\mu = \delta^\rho_\mu\,\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3\,,$$

$$\mathrm{d}x^\rho \wedge \mathrm{d}^2 S_{\mu\nu} = \delta^\rho_\nu\,\mathrm{d}^3\Sigma_\mu - \delta^\rho_\mu\,\mathrm{d}^3\Sigma_\nu\,,$$

$$\mathrm{d}x^\sigma \wedge \mathrm{d}x^\rho \wedge \mathrm{d}^2 S_{\mu\nu} = (\delta^\sigma_\mu \delta^\rho_\nu - \delta^\sigma_\nu \delta^\rho_\mu)\,\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3\,.$$

If $\mathrm{d}(h^{\mu\nu}_{ab})/\mathrm{d}x^\mu = 0$ where $\mathrm{d}/\mathrm{d}x^\mu = \partial_\mu + v^a_\mu \partial_a + v^a_{\mu\nu}\partial/\partial v^a_\nu$, $v^a_{\mu\nu} = v^a_{\nu\mu}$, then

$$h^{\mu\nu}_{ab} v^a_\mu v^b_\nu = \frac{\mathrm{d}}{\mathrm{d}x^\mu}(h^{\mu\nu}_{ab} z^a v^b_\nu)\,. \tag{6,4}$$

If $h^{\mu\nu}_{ab} = \text{const.}$ we have such a situation. In that case the momenta (6.3a) can be derived from the "equivalent" Lagrangian

$$L^* = L - \tfrac{1}{2} h^{\mu\nu}_{ab} v^a_\mu v^b_\nu\,. \tag{6,5}$$

Examples:

1. E-dynamics, with $z^\alpha = A^\alpha$, $\alpha = 0, 1, 2, 3$, $v^\alpha_\mu = \partial_\mu A^\alpha$,

$$L = -\tfrac{1}{4} F_{\mu\nu} F^{\mu\nu} - j_\alpha A^\alpha\,, \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,.$$

Possible choices of $h^{\mu\nu}_{\alpha\beta}$ are:

(i)

$$h^{\mu\nu}_{\alpha\beta} = \lambda \varepsilon^{\mu\nu}{}_{\alpha\beta}\,, \qquad \lambda = \text{const.}\,, \qquad \varepsilon_{0123} = 1 = -\varepsilon_{1023} \text{ etc.}, \tag{6,6}$$

$$\begin{aligned} h^{\mu\nu}_{\alpha\beta}\, \partial_\mu A^\alpha\, \partial_\nu A^\beta &= \lambda \varepsilon_{\mu\nu\alpha\beta} \partial^\mu A^\alpha\, \partial^\nu A^\beta \\ &= -\tfrac{1}{4}\lambda \varepsilon_{\mu\alpha\nu\beta} F^{\mu\alpha} F^{\nu\beta} = \lambda \frac{\mathrm{d}}{\mathrm{d}x^\mu}\, \varepsilon^{\mu\nu\alpha\beta} A_\alpha\, \partial_\nu A_\beta\,. \end{aligned}$$

The choice (6,6) for $h^{\mu\nu}_{\alpha\beta}$ therefore gives a "divergence" term which one encounters in the context of "anomalies" associated with U(1) axial vector currents in abelian (and nonabelian) gauge theories [Adler, 1970; Coleman, 1979].

(ii)

$$h^{\mu\nu}_{\alpha\beta} = -\lambda(\delta^\mu_\alpha \delta^\nu_\beta - \delta^\mu_\beta \delta^\nu_\alpha)\,, \tag{6,7}$$

from which we get

$$p^\mu_\alpha = -\partial^\mu A_\alpha + (1-\lambda)\,\partial_\alpha A^\mu + \lambda \delta^\mu_\alpha \partial_\nu A^\nu\,,$$

$$p^\mu_\mu = 3\lambda\, \partial_\mu A^\mu\,,$$

or, inversely,

$$\lambda(\lambda-2)\,\partial_\mu A^\alpha = g_{\mu\nu} g^{\alpha\beta} p^\nu_\beta + (1-\lambda)\, p^\alpha_\mu + (\lambda-2)\tfrac{1}{3} p^\nu_\nu\,.$$

The transformation $\partial_\mu A^\alpha \to p^\mu_\alpha$ is singular for $\lambda = 0, 2$. For $\lambda = 1$ we have

$$p^\mu_\alpha = -\partial^\mu A_\alpha + \delta^\mu_\alpha \partial_\nu A^\nu \tag{6,8}$$

with

$$p_0^0 = \partial_j A^j\,, \qquad p_j^0 = \partial_0 A^j\,.$$

Furthermore,

$$\tfrac{1}{2}h^{\mu\nu}_{\alpha\beta}\,\partial_\mu A^\alpha\,\partial_\nu A^\beta = -\tfrac{1}{2}\lambda[(\partial_\nu A^\nu)^2 - \partial_\mu A^\nu\,\partial_\nu A^\mu]$$
$$= -\tfrac{1}{2}\lambda\,\frac{\mathrm{d}}{\mathrm{d}x^\mu}(A^\mu\,\partial_\nu A^\nu - A^\nu\,\partial_\nu A^\mu)\,, \tag{6,9}$$

and the effective Lagrangian (6,5) for $\lambda = 1$ takes the form

$$L^* = -\tfrac{1}{2}\partial_\mu A^\nu\,\partial^\mu A_\nu + \tfrac{1}{2}(\partial_\nu A^\nu)^2\,. \tag{6,10}$$

The term (6,9) is well-known in the context of quantizing the e.m. field [e.g. Itzykson and Zuber, 1980, pp. 11–12].

Remarks:

(a) According to (6,8) the momentum p_0^0 is a function of the variables A^j, $j = 1, 2, 3$. If we define A^0 by Gauß's law, which yields $A^0 = -\Delta^{-1}\partial_0\partial_j A^j$ in the case of free fields, then the commutation relations $[p_j^0(x), A^k(y)]_{x^0=y^0} = -\mathrm{i}\delta_j^k\,\delta(\boldsymbol{x}-\boldsymbol{y})$, $j, k = 1, 2, 3$, are compatible with the 2-point function

$$\langle 0|A^\alpha(x)\,A^\beta(y)|0\rangle = -\mathrm{i}\Big(g^{\alpha\beta} - \frac{n^\alpha\partial^\beta + n^\beta\partial^\alpha}{n\cdot\partial}\Big)\,D^{(+)}(x-y)\,,$$
$$D^{(+)}(x) = \frac{-\mathrm{i}}{(2\pi)^3}\int \mathrm{d}^4k\;\theta(k_0)\,\delta(k^2)\,\mathrm{e}^{-\mathrm{i}k\cdot x}\,, \tag{6,11}$$

where n is a timelike unit vector, $n^2 = 1$, e.g. $n = (1, 0, 0, 0)$. In momentum space the operator acting on $D^{(+)}(x-y)$ in eq. (6,11) has the form

$$K^{\alpha\beta} = g^{\alpha\beta} - \frac{n^\alpha k^\beta + n^\beta k^\alpha}{n\cdot k}\,,$$

and $(K^{\alpha\beta})$ has the properties

$$K^{\beta\alpha} = K^{\alpha\beta}\,, \qquad k_\alpha K^{\alpha\beta} = 0 \text{ for } k^2 = 0\,, \qquad n_\alpha K^{\alpha\beta} = -k^\beta/n\cdot k\,.$$

The eigenvalues of $(K^{\alpha\beta})$ are 0, −1, −1, −2. Thus, the state space associated with the 2-point function (6,11) has a (positive) semidefinite scalar product. According to a theorem of Strocchi and Wightman [1974, Proposition 2.2] the 2-point function (6,11) is gauge equivalent to the one of the Gupta–Bleuler gauge. The propagator associated with the 2-point function (6,11) is said to correspond to a "planar gauge" [Dokshitzer et al., 1980].

The 2-point function (6,11) can be obtained from the momentum space representation

$$A_\alpha(x) = (2\pi)^{-3/2}\int \frac{\mathrm{d}^3k}{2k_0}\,(a_\alpha(k)\,\mathrm{e}^{-\mathrm{i}kx} + a_\alpha^+(k)\,\mathrm{e}^{\mathrm{i}kx})\,,$$

where [Morgenstern, 1982]

$$a_\alpha(k) = \sum_{\lambda=1}^{3} e_\alpha(\lambda)\, a(k;\lambda)\,, \qquad \alpha = 0, 1, 2, 3\,.$$

Here $e_\alpha(\lambda)$, $\lambda = 1, 2$, are the usual polarization vectors for the transversal physical photons, i.e. they have the properties [e.g., Bjorken and Drell, 1965, ch. 14]

$$e(\lambda)\cdot e(\lambda') = -\delta_{\lambda\lambda'}\,, \qquad k^\mu e_\mu(\lambda) = 0\,, \qquad \lambda = 1, 2\,,$$

$$\sum_{\lambda=1}^{2} e_\alpha(\lambda)\, e_\beta(\lambda) = -g_{\alpha\beta} + (n_\alpha k_\beta + n_\beta k_\alpha)/(n\cdot k) - k_\alpha k_\beta/(n\cdot k)^2\,.$$

With $e_\alpha(3) = k_\alpha/(n\cdot k)$ these polarization vectors, together with the commutation relations

$$[a(k;\lambda), a^+(p;\lambda')] = 2k_0\delta_{\lambda\lambda'}\delta(\boldsymbol{k}-\boldsymbol{p})\,, \qquad \lambda, \lambda' = 1, 2, 3\,,$$

yield the 2-point function (6,11) in the standard way.

Notice that with this choice of the polarization vectors we have $k^\alpha a_\alpha(k) = 0$, so that the Lorentz convention $\partial_\alpha A^\alpha = 0$ holds as an operator equation! The "states" $e_\alpha(3)\, a^+(k;3)|0\rangle$ have a vanishing norm, because $e(3)\cdot e(3) = 0$, corresponding to the fact that $e_\alpha(3)$ is a pure gauge term!

(b) The conditions $h^{\mu\nu}_{\alpha\beta} = -h^{\nu\mu}_{\alpha\beta} = -h^{\mu\nu}_{\beta\alpha}$ can be satisfied, too, by the choices

$$h^{\mu\nu}_{\alpha\beta} = \lambda F^{\mu\nu}F_{\alpha\beta}\,, \qquad \lambda = \text{const.}\,, \tag{6,12a}$$

$$h^{\mu\nu}_{\alpha\beta} = \lambda R^{\mu\nu}{}_{\alpha\beta}\,, \tag{6,12b}$$

where $R_{\mu\nu\alpha\beta}$ are the components of a Riemannian curvature tensor.

(c) As to gauge invariance we have the same situation as in section 3.4:
The 1-forms $\omega^\alpha = \mathrm{d}A^\alpha - \partial_\mu A^\alpha\, \mathrm{d}x^\mu$ are invariant under the gauge transformation

$$A^\alpha \to A^\alpha + \partial^\alpha f(x)\,, \qquad \partial_\mu A^\alpha \to \partial_\mu A^\alpha + \partial_\mu\partial^\alpha f(x)\,.$$

For that reason the form Ω is gauge invariant, if L and the coefficients $h^{\mu\nu}_{\alpha\beta}$, $h_{\alpha\beta\gamma;\mu}$ etc. are gauge invariant, too. However, after the Legendre transformation the individual coefficients H, p^μ_α etc. in general will be gauge dependent, but they will transform under gauge transformations in such a way that Ω remains invariant!

2. Another interesting example with nonvanishing constants $h^{\mu\nu}_{ab}$ is provided by the Dirac-equation [von Rieth, 1982]:
If we start with the usual Lagrangian

$$L = \frac{\mathrm{i}}{2}(\bar{z}\gamma^\mu v_\mu - \bar{v}_\mu\gamma^\mu z) - m\bar{z}z\,, \tag{6,13}$$

where z is a complex 4-component column vector and $\bar{z} = z^+\gamma^0$ a corresponding row vector. Here we have 8 field variables – because z is complex – which we take to be z and $\bar{z}$ (on the extremals we have $z = z(x) = \psi(x)$, $v_\mu = \partial_\mu \psi(x)$ etc.). As

$$\frac{\partial L}{\partial v_\mu} = \frac{\mathrm{i}}{2}\,\bar{z}\gamma^\mu\,, \qquad \frac{\partial L}{\partial \bar{v}_\mu} = -\frac{\mathrm{i}}{2}\,\gamma^\mu z\,,$$

the functional matrix $(\partial^2 L/\partial v^a_\mu\,\partial v^b_\nu)$ vanishes identically and there is no Legendre transformation.

However, since, with $\sigma^{\mu\nu} = (\mathrm{i}/2)[\gamma^\mu, \gamma^\nu]$,

$$(\bar{v}_\mu, v^{\mathrm{T}}_\mu)\begin{pmatrix} 0 & \sigma^{\mu\nu} \\ -\sigma^{\mu\nu\mathrm{T}} & 0 \end{pmatrix}\begin{pmatrix} \bar{v}^{\mathrm{T}}_\nu \\ v_\nu \end{pmatrix} = \bar{v}_\mu \sigma^{\mu\nu} v_\nu - v^{\mathrm{T}}_\mu\, \sigma^{\mu\nu\mathrm{T}} \bar{v}^{\mathrm{T}}_\nu = 2\bar{v}_\mu \sigma^{\mu\nu} v_\nu = 2\,\frac{\mathrm{d}}{\mathrm{d}x^\mu}\,(\bar{z}\sigma^{\mu\nu} v_\nu)\,,$$

we can choose the (8×8)-matrices $h^{\mu\nu} = (h^{\mu\nu}_{ab})$ to be

$$h^{\mu\nu} = \mathrm{i}\lambda\begin{pmatrix} 0 & \sigma^{\mu\nu} \\ -\sigma^{\mu\nu\mathrm{T}} & 0 \end{pmatrix}, \qquad \lambda \text{ real}\,, \tag{6,14}$$

and obtain for the Lagrangian (6,5):

$$L^* = \frac{\mathrm{i}}{2}\,(\bar{z}\gamma^\mu v_\mu - \bar{v}_\mu \gamma^\mu z) - m\bar{z}z - \mathrm{i}\lambda \bar{v}_\mu \sigma^{\mu\nu} v_\nu\,, \tag{6,15}$$

which yields the same Dirac-equations as the Lagrangian (6,13) [see e.g., Gasiorowicz, 1966, p. 90].

From this Lagrangian we obtain

$$\bar{p}^\mu := \frac{\partial L^*}{\partial v_\mu} = \frac{\mathrm{i}}{2}\,\bar{z}\gamma^\mu - \mathrm{i}\lambda \bar{v}_\nu \sigma^{\mu\nu}\,, \qquad p^\mu := \frac{\partial L^*}{\partial \bar{v}_\mu} = -\frac{\mathrm{i}}{2}\,\gamma^\mu z - \mathrm{i}\lambda \sigma^{\mu\nu} v_\nu\,. \tag{6,16}$$

Because of

$$\tau_{\mu\nu}\sigma^{\nu\lambda} = \delta^\lambda_\mu E_4 = \sigma^{\lambda\nu}\tau_{\nu\mu}\,,$$
$$\tau_{\mu\nu} = \tfrac{2}{3}\mathrm{i} g_{\mu\nu} E_4 - \tfrac{1}{3}\sigma_{\mu\nu}\,,$$

the eqs. (6,16) can be inverted:

$$\lambda v_\nu = \mathrm{i}\tau_{\nu\mu} p^\mu + \tfrac{1}{6}\mathrm{i}\gamma_\nu z\,,$$
$$\lambda \bar{v}_\nu = \mathrm{i}\bar{p}^\mu \tau_{\mu\nu} - \tfrac{1}{6}\mathrm{i}\bar{z}\gamma_\nu \tag{6,17}$$

(notice that $\gamma^\mu \tau_{\mu\nu} = \tau_{\nu\mu}\gamma^\mu = -\tfrac{1}{3}\mathrm{i}\gamma_\nu$).

We now can calculate the Hamilton function

$$H = \bar{p}^\nu v_\nu + \bar{v}_\nu p^\nu - L$$
$$= \frac{1}{\lambda}\,(\mathrm{i}\bar{p}^\nu \tau_{\nu\mu} p^\mu + \tfrac{1}{6}\mathrm{i}\bar{p}^\nu \gamma_\nu z - \tfrac{1}{6}\mathrm{i}\bar{z}\gamma_\nu p^\nu + \tfrac{1}{3}\bar{z}z) + m\bar{z}z\,, \tag{6,18}$$

from which we obtain the canonical equations

$$v_\mu = \frac{\partial H}{\partial \bar{p}^\mu} = \frac{1}{\lambda}(\mathrm{i}\tau_{\mu\nu}p^\nu + \tfrac{1}{6}\mathrm{i}\gamma_\mu z)\,,$$

$$\bar{v}_\mu = \frac{\partial H}{\partial p^\mu} = \frac{1}{\lambda}(\mathrm{i}\bar{p}^\nu\tau_{\nu\mu} - \tfrac{1}{6}\mathrm{i}\bar{z}\gamma_\mu)$$

and

$$\frac{\mathrm{d}p^\mu}{\mathrm{d}x^\mu}(x) = -\frac{\partial H}{\partial \bar{z}} = -\frac{1}{3\lambda}(-\tfrac{1}{2}\mathrm{i}\gamma_\nu p^\nu(x) + z(x)) - m\,z(x)\,,$$

$$\frac{\mathrm{d}\bar{p}^\mu}{\mathrm{d}x^\mu}(x) = -\frac{\partial H}{\partial z} = -\frac{1}{3\lambda}(\tfrac{1}{2}\mathrm{i}\bar{p}^\nu(x)\,\gamma_\nu + \bar{z}(x)) - m\,\bar{z}(x)\,, \qquad (6{,}19)$$

which are equivalent to the Dirac equations

$$\mathrm{i}\gamma^\mu\partial_\mu z(x) - m\,z(x) = 0\,, \qquad \mathrm{i}\partial_\mu\bar{z}(x)\,\gamma^\mu + m\,\bar{z}(x) = 0\,.$$

6.2. The DeDonder–Weyl canonical theory

In this theory all coefficients $h^{\mu\nu}_{ab}$ etc. vanish. The form (6,2) becomes

$$\Omega_0 = L\,\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3 + \pi^\mu_a\omega^a \wedge \mathrm{d}^3\Sigma_\mu$$
$$= a^\mu \wedge \mathrm{d}^3\Sigma_\mu - 3L\,\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3\,, \qquad (6{,}20)$$

where

$$a^\mu = L\,\mathrm{d}x^\mu + \pi^\mu_a\omega^a = -T^\mu_\nu\,\mathrm{d}x^\nu + \pi^\mu_a\,\mathrm{d}z^a\,. \qquad (6{,}21)$$

Eq. (6,20) shows that Ω_0 may be expressed by the 8 linearly independent 1-forms $\mathrm{d}x^0, \ldots, \mathrm{d}x^3$, $a^0, \ldots, a^3$ and therefore Ω_0 has the rank 8, if $n \geq 2$ (for $n = 1$ it has rank 4, because 1 p-form in $p+1$ variables always has rank p [e.g. Godbillon, 1969, p. 30]).

Replacing ω^a in Ω_0 by $\mathrm{d}z^a - v^a_\mu\,\mathrm{d}x^\mu$, we get

$$\Omega_0 = -H_{\mathrm{DW}}\,\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3 + \pi^\mu_a\,\mathrm{d}z^0 \wedge \mathrm{d}^3\Sigma_\mu\,, \qquad H_{\mathrm{DW}} = \pi^\mu_a v^a_\mu - L\,.$$

The DWHJ equation is obtained from

$$\mathrm{d}S^\mu \wedge \mathrm{d}^3\Sigma_\mu = -H_{\mathrm{DW}}\,\mathrm{d}x^0 \wedge \cdots \wedge \mathrm{d}x^3 + \pi^\mu_a\,\mathrm{d}z^a \wedge \mathrm{d}^3\Sigma_\mu\,, \qquad (6{,}22)$$

which implies

$$\partial_\mu S^\mu(x, z) + H_{\mathrm{DW}}(x, z, \pi) = 0\,, \qquad \pi^\mu_a = \partial_a S^\mu\,. \qquad (6{,}23)$$

As an application let us solve the DWHJ eq. (6,23) in the case of E-dynamics with external current $j^\alpha(x)$, provided the solution $A^\alpha(x)$ of the field equations is known [von Rieth, 1982]. In the Lorentz gauge $\partial_\mu A^\mu = 0$ the Lagrangian has the form

$$L = -\tfrac{1}{2} g_{\alpha\beta} g^{\mu\nu} v^\alpha_\mu v^\beta_\nu - j_\alpha(x)\, z^\alpha . \tag{6,24}$$

On the extremals we have $z^\alpha = A^\alpha(x)$, $v^\alpha_\mu = \partial_\mu A^\alpha(x)$, where $\Box A^\alpha(x) = j^\alpha(x)$, $\Box = \partial_0^2 - \Delta$. For $\pi^\mu_\alpha = \partial L/\partial v^\alpha_\mu$ we get

$$\pi^\mu_\alpha = -g^{\mu\nu} g_{\alpha\beta} v^\beta_\nu , \tag{6,25}$$

and therefore

$$H = v^\alpha_\mu \pi^\mu_\alpha - L = -\tfrac{1}{2} g_{\mu\nu} g^{\alpha\beta} \pi^\mu_\alpha \pi^\nu_\beta + j_\alpha(x)\, z^\alpha . \tag{6,26}$$

In order to avoid confusion, I shall use the following notation, if necessary: derivatives with respect to x^μ will be denoted by $\partial_{(\mu)}$, those with respect to z^α by ∂_α.

In order to have $\pi^j_\alpha = \partial_\alpha S^j$, $j = 1, 2, 3$, for $\pi^j_\alpha = -\partial^{(j)} A_\alpha(x)$ on the extremals $z^\alpha = A^\alpha(x)$, we make the ansatz

$$S^j(x, z) = -\partial^{(j)} A_\alpha(x)(z^\alpha - A^\alpha(x)) . \tag{6,27}$$

Then the DWHJ eq. (6,23) becomes a partial differential equation for $S^0(x, z)$:

$$\begin{aligned} &\partial_{(0)} S^0 + \hat{H}(x, z, \pi^0) = 0 , \qquad \pi^0_\alpha = \partial_\alpha S^0 , \\ &\hat{H} = -\tfrac{1}{2} g^{\alpha\beta} \pi^0_\alpha \pi^0_\beta + \tfrac{1}{2} \partial_{(j)} A^\alpha\, \partial^{(j)} A_\alpha + (\Delta A_\alpha(x))(z^\alpha - A^\alpha(x)) + j_\alpha z^\alpha . \end{aligned} \tag{6,28}$$

The characteristic equations are (derivatives with respect to $x^0 = t$ are denoted by a dot):

$$\begin{aligned} &\dot{z}^\alpha = \partial \hat{H}/\partial \pi^0_\alpha = -g^{\alpha\beta} \pi^0_\beta , \\ &\dot{\pi}^0_\alpha = -\partial \hat{H}/\partial z^\alpha = -\Delta A_\alpha(x) - j_\alpha(x) = -\partial_t^2 A_\alpha(x) , \end{aligned}$$

with the (special) solutions

$$z^\alpha(t) = A^\alpha(x) + b^\alpha , \qquad \dot{b}^\alpha = 0 , \qquad \pi^0_\alpha = -\partial_t A_\alpha(x) .$$

For $\dot{S}^0$ we therefore get the equation

$$\dot{S}_0 = -\hat{H} + \dot{z}^\alpha\, \pi^0_\alpha(t) = -\tfrac{1}{2} \partial_{(\mu)} A^\alpha\, \partial^{(\mu)} A_\alpha - b^\alpha \Delta A_\alpha - j_\alpha (A^\alpha + b^\alpha) ,$$

which, with the help of the field equations, yields

$$S^0(x^0; b, \boldsymbol{x}) = \text{const.}(\boldsymbol{x}) + \int_0^{x^0} L[A(x), \partial A(x)]\, \mathrm{d}t - b^\alpha [\partial_0 A_\alpha]_0^{x^0} . \tag{6,29}$$

Taking the "const." to be $-b^\alpha(\partial_t A_\alpha)_{t=0}$ and replacing b^α by $z^\alpha - A^\alpha(x)$ we finally get the solution

$$S^0(x,z) = -(z^\alpha - A^\alpha(x))\,\partial_{(0)}A_\alpha(x) + \int_0^{x^0} L[A(x), \partial A(x)]\,\mathrm{d}t\,, \tag{6,30}$$

which has the right transversality property $\pi^0_\alpha = \partial_\alpha S^0 = -\partial_{(0)}A_\alpha$ and which obeys the integrability conditions

$$\partial_{(\nu)}\varphi^\alpha_\mu + \partial_\beta\varphi^\alpha_\mu\varphi^\beta_\nu = \partial_{(\mu)}\varphi^\alpha_\nu + \partial_\beta\varphi^\alpha_\nu\varphi^\beta_\mu\,,$$

$$\varphi^\alpha_\mu = \begin{cases} -g^{\alpha\beta}\partial_\beta S^0 & \text{for } \mu = 0 \\ g^{\alpha\beta}\partial_\beta S^j & \text{for } \mu = j = 1,2,3\,. \end{cases}$$

Let $S^\mu(x,z;a)$ be a solution of the eq. (6,23) depending on a parameter a. If $z^b = f^b(x)$ is an extremal for which $\pi^\mu_b(x) = \partial_b S^\mu(x, z = f(x))$, then, in the same way as in section 4.2, one can show that the functions

$$G^\mu(x) = \frac{\partial S^\mu}{\partial a}(x, z = f(x); a) \tag{6,31}$$

are the components of a conserved current!
Example [von Rieth, 1982]:
Suppose the vector potential $A^\alpha(x;a)$ is a solution of Maxwell's equations with external current $j^\alpha(x)$ and that $A^\alpha(x;a)$ depends on a parameter a (as a consequence of boundary conditions, for instance), which does not occur in $j^\alpha(x)$. Then we get from eqs. (6,29) and (6,27)

$$\begin{aligned} G^0(x) &= \left.\frac{\partial S^0}{\partial a}\right|_{z^\alpha = A^\alpha(x)} = \partial_0 A_\alpha(x)\frac{\partial}{\partial a}A^\alpha(x;a) + \int_0^{x^0} \mathrm{d}t\,\frac{\partial L}{\partial a}[A(x;a), \partial A(x;a)]\,, \\ G^j(x) &= \left.\frac{\partial S^j}{\partial a}\right|_{z^\alpha = A^\alpha(x)} = \partial^j A_\alpha(x)\frac{\partial}{\partial a}A^\alpha(x;a)\,. \end{aligned} \tag{6,32}$$

Using the field equations one verifies easily that indeed $\partial_\mu G^\mu(x) = 0$.
Application: Suppose we have $j^\alpha(x) = 0$, then the plane waves with

$$A^0 = 0,\ A^j = f^j(t - \boldsymbol{n}\cdot\boldsymbol{x})\,, \qquad \boldsymbol{n}\cdot\boldsymbol{f} = 0,\ \boldsymbol{n}^2 = 1\,,$$

$$\boldsymbol{n} = (n^1, n^2, n^3) = (\cos\alpha\sin\beta, \sin\alpha\sin\beta, \cos\beta)$$

are solutions of the homogeneous Maxwell equations depending on the parameters α and β. We have

$$\partial_0 A^j = \dot f^j\,, \qquad \partial_k A^j = -n^k \dot f^j\,,$$

so that $L = -\frac{1}{2}\partial_\mu A^j\,\partial^\mu A_j = 0$.

Since $\partial_0 A^j = \dot{f}^j = -E^j$ and, for instance,

$$\partial A^j / \partial \alpha = \dot{f}^j \sin\beta\,(x^1 \sin\alpha - x^2 \cos\alpha)\,,$$

we get for the current components (6,25)

$$\begin{aligned} G^0(x) &= -\boldsymbol{E}^2(x)\sin\beta\,(x^1\sin\alpha - x^2\cos\alpha)\,, \\ G^j(x) &= -n^j\,\boldsymbol{E}^2(x)\sin\beta\,(x^1\sin\alpha - x^2\cos\alpha)\,. \end{aligned} \tag{6,33}$$

Since for plane waves the energy continuity equation $\partial_0(\boldsymbol{E}^2) + \operatorname{div}(\boldsymbol{n}\boldsymbol{E}^2) = 0$ holds and since $\operatorname{div}[\boldsymbol{n}\,(x^1\sin\alpha - x^2\cos\alpha)] = 0$, the current (6,33) is indeed conserved. The structure of that current is of the following type: Suppose $g^\mu(x)$ is a conserved current, $\partial_\mu g^\mu = 0$. If $f(\boldsymbol{x})$ does not depend on x^0 and if $g^j \partial_j f(\boldsymbol{x}) = 0$, then $G^\mu(x) = f(\boldsymbol{x})\,g^\mu(x)$ is a conserved current, too!

All other concepts of interest associated with the DWHJ theory, like that of a complete integral etc., are the same as in the case of systems with 2 independent variables, discussed in chapter 4 and will not be repeated here.

6.3. Carathéodory's canonical theory

With the definitions

$$\begin{aligned} a^\mu &= L\,\mathrm{d}x^\mu + \pi^\mu_a \omega^a = -T^\mu_\nu\,\mathrm{d}x^\nu + \pi^\mu_a\,\mathrm{d}z^a\,, \\ \theta^\mu &= -H\,\mathrm{d}x^\mu + p^\mu_a\,\mathrm{d}z^a\,, \qquad T^\mu_\nu = \pi^\mu_a v^a_\nu - \delta^\mu_\nu L\,, \qquad \mu,\nu = 1,\dots,m\,, \end{aligned}$$

Carathéodory's canonical framework is defined by the following expression for the fundamental canonical form Ω_c:

$$\Omega_c = L^{1-m}\,a^1 \wedge \cdots \wedge a^m = (-H_c)^{1-m}\,\theta^1 \wedge \cdots \wedge \theta^m\,. \tag{6,34}$$

Expressing a^μ in the basis $(\mathrm{d}x^\mu, \mathrm{d}z^a)$ and comparing both sides of eq. (6,34) we obtain

$$\begin{aligned} H_c &= (-L)^{1-m}\det(T^\mu_\nu)\,, \\ p^\mu_a &= (-L)^{1-m}\,\bar{T}^\mu_\rho \pi^\rho_a\,, \qquad \text{or} \qquad H_c\pi^\mu_a = T^\mu_\rho p^\rho_a\,. \end{aligned} \tag{6,35}$$

Again we see that for "large" L

$$H_c = H_{\mathrm{DW}} + \mathrm{O}(1/L)\,, \qquad H_{\mathrm{DW}} = \pi^\mu_a v^a_\mu - L\,,$$

$$p^\mu_a = \pi^\mu_a + \mathrm{O}(1/L)\,,$$

which shows that the DW canonical framework – which is the conventional one in physics – is that approximation of Carathéodory's canonical theory which neglects all invariants of the matrix $(\pi^\mu_a v^a_\nu)$ but its trace!

If we define $R = (v^a_\mu \pi^\mu_b - \delta^a_b L) = v \cdot \pi - LE_n$, we have, as in section 5.1, with the help of formula (2,40a)

$$|T| = |-LE_m + \pi v| = \left| \begin{pmatrix} E_n & v \\ -\pi & -LE_m \end{pmatrix} \right| = (-L)^m \left| E_n - \frac{1}{L} v \cdot \pi \right| = (-L)^{m-n} |R| , \tag{6,36}$$

and therefore

$$H_c = (-L)^{1-n} |R| ,$$

like eq. (5,18). The eqs. (5,26) hold for arbitrary m. The same is true for the eqs. (5,27) and (5,29), and so on!

The CHJ equation is obtained from the relation

$$\mathrm{d}S^1 \wedge \cdots \wedge \mathrm{d}S^m(x, z) = (-H_c)^{1-m} \theta^1 \wedge \cdots \wedge \theta^m = L^{1-m} a^1 \wedge \cdots \wedge a^m , \tag{6,37}$$

which gives

$$|(\partial_\mu S^\nu)| + H_c(x, z, p) = 0 , \tag{6,38a}$$

where

$$p^\mu_a = (\overline{\partial S})^\mu_\rho \, \partial_a S^\rho ,$$

or

$$p^\rho_a \, \partial_\rho S^\mu = |(\partial_\rho S^\nu)| \, \partial_a S^\mu = -H_c \, \partial_a S^\mu , \tag{6,38b}$$

and

$$L = |\Delta| , \qquad \Delta = (\Delta^\mu_\nu = \partial_\nu S^\mu + v^a_\nu \, \partial_a S^\mu) , \tag{6,39a}$$

$$\pi^\mu_a = (\bar{\Delta})^\mu_\rho \, \partial_a S^\rho = \partial |\Delta| / \partial v^a_\mu . \tag{6,39b}$$

If $S^\mu(x, z; a)$ is a solution of the CHJ equation which depends on a parameter a and if $z^b = f^b(x)$ is an extremal for which the relation

$$p^\mu_b(x) = (\overline{\partial S})^\mu_\rho \, \partial_b S^\rho(x, z = f(x); a)$$

holds, then the current

$$G^\mu(x) = \frac{\partial S^\rho}{\partial a} \bar{\Delta}^\mu_\rho(x, z = f(x); a) \tag{6,40}$$

is conserved. The complete integral of the CHJ eq. (6,38) is defined as in section 5.6. The canonical E. Hölder transformation

$$x^1 \to \hat{x}^1 = x^1, \qquad x^{\bar{\mu}} \to \hat{x}^{\bar{\mu}} = \sigma^{\bar{\mu}} = S^{\bar{\mu}}(x, z), \qquad \bar{\mu} = 2, \ldots, m,$$
$$z^a \to \hat{z}^a = z^a \tag{6,41}$$

again transforms Ω_c into

$$\Omega_c = (-\hat{H}\, d\hat{x}^1 - \hat{p}^1_a\, d\hat{z}^a) \wedge d\sigma^2 \wedge \cdots \wedge d\sigma^m, \tag{6,42}$$

so that on the surfaces $\sigma^{\bar{\mu}}$ = const. the canonical field equations are the same as in mechanics. All the details can be derived in complete analogy to the discussion in section 5.7.

Bibliographical notes

The references here are the same as those in chapters 3–5.

7. Hamilton–Jacobi theories for systems with *m*-dimensional action integrals which are invariant under reparametrization

7.1. Consequences of the homogeneity properties of the Lagrangian function

Let $\Sigma_0^{(m)} = \{x(t) \in \mathrm{R}^n,\ t \in \mathrm{G}^m \subset \mathrm{R}^m,\ m < n\}$ be an m-dimensional hypersurface in an n-dimensional Euclidean space. The surface $\Sigma_0^{(m)}$ is supposed to be an extremal of the action integral

$$A = \int_{\Sigma^{(m)}} L(x, u)\, dt^1 \cdots dt^m, \tag{7,1}$$

where the variables u^i_μ, $\mu = 1, \ldots, m$, $i = 1, \ldots, n$ become the derivatives $\partial_{(\mu)} x^i = \partial x^i/\partial t^\mu$ on any surface $\Sigma^{(m)}$ in the neighborhood of $\Sigma_0^{(m)}$. In order that the surface $\Sigma_0^{(m)}$ has an intrinsic geometrical meaning the action integral A has to be invariant under any smooth one-to-one reparametrization

$$t^\mu \to \hat{t}^\mu = \hat{T}^\mu(t), \qquad \hat{t}^\mu \to t^\mu = T^\mu(\hat{t}), \qquad |(\partial \hat{T}^\mu/\partial t^\nu)| > 0,$$

of the surfaces $\Sigma^{(m)}$. As

$$\hat{u}^i_\mu = \partial x^i/\partial \hat{t}^\mu = \partial_{(\nu)} x^i\, \partial T^\nu/\partial \hat{t}^\mu = u^i_\nu\, \partial T^\nu/\partial \hat{t}^\mu,$$
$$d\hat{t}^1 \wedge \cdots \wedge d\hat{t}^m = |(\partial \hat{T}/\partial t)|\, dt^1 \wedge \cdots \wedge dt^m,$$

it follows from

$$L(x, \hat{u})\, d\hat{t}^1 \wedge \cdots \wedge d\hat{t}^m = L(x, u)\, dt^1 \wedge \cdots \wedge dt^m$$

that

$$L[x, u\cdot(\partial T/\partial\hat{t})] = |(\partial T/\partial\hat{t})|\, L(x, u)\,. \tag{7,2}$$

Physically interesting examples of systems the Lagrangians of which have the property (7,2) are

(i) $m = 1$: geodesics, where $L(x, u) = (g_{ij}(x)u^i u^j)^{1/2}$. (7,3a)
(ii) $m = 2$, $n = 4$: relativistic strings [Nambu, 1970; Rebbi, 1974; Scherk, 1975], where

$$L(x, u) = \text{const.}\,[(u_1\cdot u_2)^2 - (u_1)^2(u_2)^2]^{1/2}\,, \tag{7,3b}$$
$$a\cdot b = a^0 b^0 - \boldsymbol{a}\cdot\boldsymbol{b}\,.$$

In the following we choose units such that the constant in eq. (7,3b) takes the value 1.
(iii) $m = 3$, $n = 4$: relativistic vibrating membranes [Collins and Tucker, 1976] with

$$L(x, u) = (\det(h_{\mu\nu}))^{1/2}\,, \qquad h_{\mu\nu} = u_\mu\cdot u_\nu, \qquad \mu, \nu = 1, 2, 3\,. \tag{7,3c}$$

The homogeneity condition (7,2) has the following implications [Wilkins Jr., 1944]:

$$u^i_\lambda\, \partial L/\partial u^i_\mu = \delta^\mu_\lambda L\,, \tag{7,4a}$$

$$u^i_\lambda \frac{\partial^2 L}{\partial u^i_\mu\, \partial u^j_\nu} = \delta^\mu_\lambda \frac{\partial L}{\partial u^j_\nu} - \delta^\nu_\lambda \frac{\partial L}{\partial u^j_\mu}\,, \tag{7,4b}$$

or

$$u^i_\lambda\left(\frac{\partial^2 L}{\partial u^i_\mu\, \partial u^j_\nu} + \frac{\partial^2 L}{\partial u^i_\nu\, \partial u^j_\mu}\right) = 0\,. \tag{7,4c}$$

The eqs. (7,4a) are obtained from eq. (7,2) by differentiating with respect to $\tau^\mu_\nu := \partial T^\mu/\partial\hat{t}^\nu$ and setting (τ^μ_ν) equal to the unit matrix E_m afterwards. Differentiating the relations (7,4a) with respect to the variable u^j_ν gives the eqs. (7,4b).

Another important consequence of the homogeneity condition (7,2) is that it is equivalent to $L(x, u)$ being a homogeneous function of the Grassmann coordinates

$$v^{i_1\cdots i_m} := \left|\begin{pmatrix} u^{i_1}_1 & \cdots & u^{i_1}_m \\ \vdots & & \vdots \\ u^{i_m}_1 & & u^{i_m}_m \end{pmatrix}\right|\,, \qquad i_\mu = 1, \ldots, n\,, \tag{7,5}$$

alone [Frechét, 1905; Gross, 1916], that is we have

$$L = L(x, v)\,, \qquad L(x, kv) = k\, L(x, v)\,, \qquad k > 0\,. \tag{7,6}$$

As the coordinates $v^{i_1\cdots i_m}$ are completely antisymmetric in the indices $i_1, \ldots, i_m$, we have $\binom{n}{m}$ Grassmann coordinates. In the following we shall denote by $v^{(i)}$ the so-called "strict" components $v^{(i_1\cdots i_m)} = v^{i_1\cdots i_m}$,

$i_1 < i_2 < \cdots < i_m$. For $2 \leq m \leq n-2$ these coordinates are not independent, but obey the bilinear relations [Hodge and Pedoe II, 1952, ch. 14]

$$\sum_{\mu=1}^{m+1} (-1)^{\mu}\, v^{i_1\ldots i_{m-1}j_\mu} v^{j_1\ldots j_{\mu-1}j_{\mu+1}\ldots j_{m+1}} = 0\,, \qquad i_1 < \cdots < i_{m-1},\, j_1 < \cdots < j_{m+1}\,, \tag{7,7}$$

which define a Grassmann manifold. Suppose the number of the relations (7,7) is r. Dividing them by $N := (v^{(i)}v^{(i)})^{1/2}$ and denoting the resulting l.h. side by $g_\rho(v)$, $\rho = 1, \ldots, r$, we have

$$v^{(i)} \frac{\partial}{\partial v^{(i)}} g_\rho = g_\rho\,.$$

Example:
For the relativistic string (7,3b) the number r of constraints (7,7) is one:

$$\begin{aligned} g(v) &= (v^{01}v^{23} - v^{02}v^{13} + v^{03}v^{12})N^{-1} = 0\,,\\ N &= +[(v^{01})^2 + \cdots + (v^{23})^2]^{1/2}\,. \end{aligned} \tag{7,8}$$

It follows from $L(x, kv) = k\,L(x, v)$ that

$$v^{(i)} \frac{\partial L}{\partial v^{(i)}} = \sum_{i_1<\cdots<i_m} v^{i_1\ldots i_m} \frac{\partial L}{\partial v^{i_1\ldots i_m}} = L \tag{7,9a}$$

and

$$v^{(i)} \frac{\partial^2 L}{\partial v^{(i)}\,\partial v^{(k)}} = 0\,. \tag{7,9b}$$

The last equation implies

$$\left| \left(\frac{\partial^2 L}{\partial v^{(i)}\,\partial v^{(k)}} \right) \right| = 0\,. \tag{7,10}$$

The vanishing of the determinant (7,10) means that we are dealing with a "singular" variational problem. For $m = 1$, where $v^{(i)} = u^i = \dot{x}^i$, such systems have been treated mathematically by Carathéodory [1935, ch. 13] and, from a physical point of view, by Dirac [1933, 1964] and others. The case $m \geq 2$ more recently was analyzed – according to the ideas of Carathéodory – by Velte [1953, 1954], Klötzler [1961], Fuchs [1973], Beckert [1974].

For $m = 1$ Carathéodory [1935, ch. 13] showed that the usual regularity condition $|(\partial^2 L/\partial v^i \partial v^j)| \neq 0$ in the case of "parametric" problems can be replaced by

$$\left| \begin{pmatrix} \dfrac{\partial^2 L}{\partial u^i \partial u^k} & \begin{matrix} u^1 \\ \vdots \\ u^n \end{matrix} \\ u^1 \ldots u^n & 0 \end{pmatrix} \right| \neq 0 \tag{7,11}$$

if the rank of $(\partial^2 L/\partial u^i\,\partial u^k)$ is $n-1$. For details we refer to Carathéodory's monograph. [See also Hestenes, 1966, ch. 2, §8.]

In the cases $m \geq 2$ we have to take the constraints (7,7) into account. This can be done by means of Lagrangian multipliers:

$$L \to \hat{L} = L(x, v) + \lambda^\rho g_\rho(v)\,.$$

According to Velte [1954] a generalization of the regularity condition (7,11) here takes the form

$$D := \left| \begin{pmatrix} \dfrac{\partial^2(\frac{1}{2}L^2)}{\partial v^{(i)}\,\partial v^{(j)}} & \dfrac{\partial g_\rho}{\partial v^{(i)}} \\ \partial g_\rho/\partial v^{(j)} & 0_r \end{pmatrix} \right| \neq 0\,, \tag{7,12}$$

where 0_r is the $(r \times r)$ zero matrix!

Example:

For the string Lagrangian (7,3b) with the constraint (7,8) the determinant (7,12) has the value $-L^2/N^2$ on $g(v) = 0$ (for the calculation the formula (2,40b) is useful), that is to say, it is "regular" if $L \neq 0$!

7.2. Legendre transformation

The regularity of the determinant (7,12) is important in the context of the following Legendre transformation, too: This transformation from the set of variables $v^{(i)}$, λ^ρ onto the set $p_{(i)}$, μ_ρ, $\rho = 1, \ldots, r$, is defined by

$$p_{(i)}(v) = \hat{L}\,\partial\hat{L}/\partial v^{(i)}\,, \qquad \mu_\rho = g_\rho(v)\,. \tag{7,13}$$

Its functional determinant

$$\left| \begin{pmatrix} \dfrac{\partial^2(\frac{1}{2}\hat{L}^2)}{\partial v^{(i)}\,\partial v^{(j)}} & \hat{L}\dfrac{\partial g_\rho}{\partial v^{(i)}} + g_\rho \dfrac{\partial \hat{L}}{\partial v^{(i)}} \\ \partial g_\rho/\partial v^{(j)} & 0_r \end{pmatrix} \right|$$

for $\lambda^\rho = 0$, $g_\rho(v) = 0$ takes the value DL^r. If $DL \neq 0$, then we can solve eqs. (7,13) for the variables

$$v^{(i)} = V^{(i)}(x, p, \mu)\,, \qquad \lambda^\rho = \Lambda^\rho(x, p, \mu)\,.$$

If one defines

$$G(x, p, \mu) := \hat{L}[x, v = V(x, p, \mu), \lambda = \Lambda(x, p, \mu)]\,, \tag{7,14a}$$

$$G^{(i)}(x, p, \mu) := V^{(i)}/G \tag{7,14b}$$

then it follows from $v^{(i)}p_{(i)} = \hat{L}^2 = \hat{L}G$ that

$$\mathrm{d}G - G^{(i)}\,\mathrm{d}p_{(i)} = -(\mathrm{d}\hat{L} - \hat{L}_{(i)}\,\mathrm{d}v^{(i)})\,, \qquad \hat{L}_{(i)} := \partial\hat{L}/\partial v^{(i)}\,, \tag{7,15}$$

or

$$\partial_i G\, dx^i + \left(\frac{\partial G}{\partial p_{(i)}} - G^{(i)}\right) dp_{(i)} + \frac{\partial G}{\partial \mu_\rho} d\mu_\rho = -\partial_i \hat{L}\, dx^i - g_\rho\, d\lambda^\rho , \qquad \partial_i G = \partial G/\partial x^i , \tag{7,16}$$

which implies that for $g_\rho(v) = 0$ and $d\mu_\rho = 0$

$$\partial_i G = -\partial_i \hat{L} , \qquad \partial G/\partial p_{(i)} = G^{(i)} . \tag{7,17}$$

If we define $H(x, p) := G(x, p, \mu = 0)$, we get for $\lambda^\rho = 0$, $g_\rho(v) = 0$, $\hat{L} \neq 0$:

$$p_{(i)} = L\, L_{(i)} , \qquad H = L , \qquad H(x, kp) = k\, H(x, p) ,$$
$$\partial H/\partial p_{(i)} = v_{(i)}/L , \qquad \partial_i H = -\partial_i L . \tag{7,18}$$

Example:
From the Lagrangian (7,3b) of the relativistic string we get [Rinke, 1980]

$$L_{(\alpha\beta)} := \partial L/\partial v^{(\alpha\beta)} = -v_{(\alpha\beta)}/L(v) , \qquad L(v) = (-v_{(\alpha\beta)} v^{(\alpha\beta)})^{1/2} , \qquad v_{(\alpha\beta)} = g_{\alpha\gamma} g_{\beta\delta} v^{(\gamma\delta)} ,$$

and therefore

$$p_{(\alpha\beta)} = -v_{(\alpha\beta)} , \qquad \alpha, \beta = 0, \dots, 3 , \tag{7,19a}$$

$$H(p) = (-p_{(\alpha\beta)} p^{(\alpha\beta)})^{1/2} . \tag{7,19b}$$

Up to now we have ignored the following problem: It is by no means obvious that the derivatives $L_{(i)} := \partial L/\partial v^{(i)}$ belong to the Grassmann manifold defined by the constraints $g_\rho(v) = 0$, $\rho = 1, \dots, r$, too, i.e. it is a priori not clear that the $g_\rho(p) = 0$, $\rho = 1, \dots, r$. However, it was shown by H. Kneser [1940], Velte [1954] and Barthel [1958] that it is always possible, for $L \neq 0$, to replace that Lagrangian L by an equivalent one, $L^* = L + M(x, v)$, where M vanishes for $g_\rho(v) = 0$, $\rho = 1, \dots, r$, such that $g_\rho(L^*_{(i)}) = 0$, $\rho = 1, \dots, r$.

7.3. Hamilton–Jacobi theories

A HJ theory for systems invariant under reparametrizations is obtained as follows [Velte 1954]: Consider the m-form

$$\omega = L_{(i_1 \dots i_m)}\, dx^{i_1} \wedge \cdots \wedge dx^{i_m} , \tag{7,20}$$

which for $x^i = x^i(t)$ becomes

$$\omega = L_{(i)} v^{(i)}\, dt^1 \wedge \cdots \wedge dt^m = L\, dt^1 \wedge \cdots \wedge dt^m .$$

Suppose that for a family of extremals $x(t)$ there exist functions $\varphi^{(i)}(x)$ such that $v^{(i)}(t) = \varphi^{(i)}(x(t))$ and

$$d\tilde{\omega} = d[L_{(i_1 \dots i_m)}(x, v = \varphi(x))\, dx^{i_1} \wedge \cdots \wedge dx^{i_m}] = 0\,, \tag{7,21}$$

then the set $\{(x, \varphi(x))\}$ is said to define a "geodesic field" for that family of extremals.

Because the coefficients $L_{(i)}$ obey the relations (7,7) the m-form (7,20) is decomposable [Satake, 1975, p. 258] and therefore has rank m. Thus, according to Poincaré's lemma the property (7,21) implies that locally

$$\tilde{\omega} = dS^1 \wedge \cdots \wedge dS^m(x)\,, \tag{7,22}$$

so that

$$L_{(i)}(x, v = \varphi(x)) = s_{(i)}(x)\,, \qquad s_{i_1 \dots i_m} = \left| \begin{pmatrix} \partial_{i_1} S^1 & \dots & \partial_{i_1} S^m \\ \vdots & & \vdots \\ \partial_{i_m} S^1 & \dots & \partial_{i_m} S^m \end{pmatrix} \right| \,. \tag{7,23}$$

According to eqs. (7,18) we have $H(x, p_{(i)} = L_{(i)}L) = L$ therefore $H(x, L_{(i)}) = 1$, i.e. the functions $s_{(i)}(x)$ have to obey the HJ equation

$$H(x, s_{(i)}) = 1\,, \tag{7,24}$$

which again is one partial differential equation for m functions $S^1(x), \dots, S^m(x)$. Therefore $m-1$ of them can be chosen appropriately, in accordance with the "transversality" conditions (7,23)!
Example:
For the relativistic string (7,3b) the eqs. (7,19) imply the HJ equation

$$H = (-s_{(\alpha\beta)} s^{(\alpha\beta)})^{1/2} = 1\,, \qquad s_{\alpha\beta}(x) = \partial_\alpha S^1\, \partial_\beta S^2 - \partial_\alpha S^2\, \partial_\beta S^1\,. \tag{7,25}$$

If $S^\mu(x)$ is a solution of the HJ eq. (7,24) and if $x^i = f^i(t)$ is a solution of the first order differential equations

$$L_{(i)}(x(t), v(t)) = s_{(i)}(x(t))\,, \tag{7,26}$$

then the functions $x^i = f^i(t)$ are solutions of the Euler–Lagrange equations

$$\partial_{(\mu)}\, \partial L / \partial u^i_\mu - \partial_i L = 0\,,$$

too [Klötzler, 1961].
Proof: Expanding the determinant $|\Delta|$ of the matrix

$$\Delta = (\Delta^\mu_\nu = \partial_i S^\mu u^i_\nu)$$

with respect to the minors $s_{(i)}$ gives [Satake, 1975, p. 78]:

$$|\Delta| = s_{(i)} v^{(i)}\,. \tag{7,27}$$

Differentiating eq. (7,27) with respect to u^j_μ and $\partial_j S^\mu$ respectively one obtains, see eq. (5,24),

$$s_{(i)}\,\partial v^{(i)}/\partial u^j_\mu = \bar{\Delta}^\mu_\nu\,\partial_j S^\nu\,, \tag{7,28a}$$

$$v^{(i)}\,\partial s_{(i)}/\partial(\partial_j S^\mu) = \bar{\Delta}^\nu_\mu\,u^j_\nu\,. \tag{7,28b}$$

Therefore we have

$$\frac{\partial L}{\partial u^j_\mu} = \frac{\partial L}{\partial v^{(i)}}\frac{\partial v^{(i)}}{\partial u^j_\mu} = s_{(i)}\frac{\partial v^{(i)}}{\partial u^j_\mu} = \bar{\Delta}^\mu_\nu\,\partial_j S^\nu\,, \tag{7,29}$$

and, because $\mathrm{d}(\bar{\Delta}^\mu_\nu)/\mathrm{d}t^\mu = 0$, eq. (5,76),

$$\frac{\mathrm{d}}{\mathrm{d}t^\mu}\left(\frac{\partial L}{\partial u^i_\mu}\right) = \bar{\Delta}^\mu_\nu\,\frac{\mathrm{d}}{\mathrm{d}t^\mu}(\partial_i S^\nu) = \partial_j\partial_i S^\nu u^j_\mu \bar{\Delta}^\mu_\nu\,. \tag{7,30}$$

Using eqs. (7,28b) we get

$$\partial_i\partial_j S^\nu u^j_\mu \bar{\Delta}^\mu_\nu = \partial_i\partial_j S^\nu v^{(k)}\,\partial s_{(k)}/\partial(\partial_j S^\nu) = v^{(j)}\,\partial_i s_{(j)}\,,$$

and since

$$\partial_j s_{(i)} = \mathrm{D}_j L_{(i)}(x, v = \varphi(x)) = \partial_j L_{(i)} + \frac{\partial^2 L}{\partial v^{(k)}\,\partial v^{(i)}}\,\partial_j\varphi^{(k)}\,,$$

the eqs. (7,9) imply

$$v^{(j)}\partial_i s_{(j)} = \partial_i L\,, \tag{7,31}$$

q.e.d.

As the functions $S^\mu(x)$ do not depend on the variables t, the surfaces $S^\mu(x) = \sigma^\mu = \text{const.}$, $\mu = 1, \ldots, m$, are $(n-m)$-dimensional. The geometrical interpretation of the transversality properties of extremals and wave fronts in this case are as follows [Velte, 1953, 1954]:

Suppose we have m linearly independent ("normal") vectors $h^{(\mu)} = (h^{(\mu)}_1, \ldots, h^{(\mu)}_n)$, $\mu = 1, \ldots, m$, which span an m-dimensional subspace and $k_{(\hat{\mu})} = (k^1_{(\hat{\mu})}, \ldots, k^n_{(\hat{\mu})})$, $\hat{\mu} = m+1, \ldots, n$, $(n-m)$ linearly independent (tangent) vectors, "orthogonal" to all $h^{(\mu)}$, i.e. we have

$$h^{(\mu)}_i\,k^i_{(\hat{\mu})} = 0\,, \qquad \mu = 1, \ldots, m, \hat{\mu} = m+1, \ldots, n\,. \tag{7,32}$$

In our applications the m vectors $h^{(\mu)}$ will be the gradients $(\partial_1 S^\mu, \ldots, \partial_n S^\mu)$. Consider now the two sets of Grassmann coordinates

$$h_{(i)} = h_{(i_1\cdots i_m)} = \left|\begin{pmatrix} h^{(1)}_{i_1} \cdots h^{(m)}_{i_1} \\ \vdots \qquad \vdots \\ h^{(1)}_{i_m} \cdots h^{(m)}_{i_m} \end{pmatrix}\right|, \qquad i_\mu = 1, \ldots, n\,, \tag{7,33a}$$

– if $h^{(\mu)} = (\partial_1 S^\mu, \ldots, \partial_n S^\mu)$, then we have $h_{(i)} = s_{(i)}$ –, and

$$k^{(\hat{i})} = k^{(i_{m+1}\cdots i_n)} = \left|\begin{pmatrix} k^{i_{m+1}}_{(m+1)} & \cdots & k^{i_{m+1}}_{(n)} \\ \vdots & & \vdots \\ k^{i_n}_{(m+1)} & \cdots & k^{i_n}_{(n)} \end{pmatrix}\right|, \qquad i_{\bar{\mu}} = 1, \ldots, n. \tag{7,33b}$$

Because of the relations (7,32) the coordinates (7,33a) and (7,33b) are not independent: If $(\hat{i})$ is the sequence $(i_{m+1}, \ldots, i_n)$ of indices obtained from $1, \ldots, n$ by deleting the set $(i) = (i_1, \ldots, i_m)$, then [Hodge and Pedoe I, 1953, ch. VII]:

$$k^{(\hat{i})} = \rho(-1)^s h_{(i)}, \qquad s = \sum_{\mu=1}^{m} i_\mu + \tfrac{1}{2}m(n+1), \qquad \rho = \text{const.}. \tag{7,34}$$

Given a Lagrangian $L(x, v)$ then one can associate with each m-dimensional surface element (x, v) an $(n-m)$-dimensional transversal surface element

$$k^{(\hat{i})} = (-1)^s L_{(i)}(x, v). \tag{7,35}$$

If $u_\mu = (u^1_\mu, \ldots, u^n_\mu)$, $\mu = 1, \ldots, m$, are m linearly independent tangent vectors of the extremal $x^i = f^i(t)$, then we have

$$\det(u_1, \ldots, u_m, k_{(m+1)}, \ldots, k_{(n)}) = v^{(i)} L_{(i)} = L, \tag{7,36}$$

so that the n vectors $u_1, \ldots, k_{(n)}$ are linearly independent, if $L \neq 0$.

If $L \neq 0$, then one can always choose a parametrization such that $L = 1$ on the extremal. For instance, if $L(x(t), v(t)) = \Lambda(t) \neq 1$, then

$$\hat{t}^1 = \int_0^{t^1} \Lambda(t)\, dt^1, \qquad \hat{t}^{\bar{\mu}} = t^{\bar{\mu}}, \quad \bar{\mu} = 2, \ldots, m,$$

is such a parametrization.

An $(n-m)$-dimensional plane through the point $x = f(t)$, t fixed, of an extremal is given by

$$y^i(t, \xi) = f^i(t) + k^i_{(\hat{\mu})}\xi^{\hat{\mu}}, \tag{7,37}$$

where the vectors $k_{(\hat{\mu})}(t)$ determine the direction and the variables $\xi^{\hat{\mu}}$, $\hat{\mu} = m+1, \ldots, n$, the points of the planes. The functional determinant

$$\partial(y^1, \ldots, y^n)/\partial(t^1, \ldots, t^m, \xi^{m+1}, \ldots, \xi^n)$$

is the same as in eq. (7,36) and it is therefore possible to solve the eqs. (7,37) in a neighborhood of $x = f(t)$ for t^μ and $\xi^{\hat{\mu}}$:

$$t^\mu = \tilde{S}^\mu(x) \tag{7,38a}$$

$$\xi^{\hat{\mu}} = \xi^{\hat{\mu}}(x), \tag{7,38b}$$

and the $(n-m)$-dimensional hyperplanes (7,37) through $f(t)$ can be characterized by the equations $\tilde{S}^\mu(x) = t^\mu = \text{const}$. The m vectors $(\partial_1\tilde{S}^\mu, \ldots, \partial_n\tilde{S}^\mu)$, $\mu = 1, \ldots, m$, span the space dual to the $(n-m)$-dimensional planes (7,37) and therefore the Grassmann coordinates $\tilde{s}_{(i)}$ must be proportional to $L_{(i)}$. Furthermore, since $t^\mu = \tilde{S}^\mu(x(t))$, we have

$$\tilde{\Delta}^\mu_\nu = \mathrm{d}\tilde{S}^\mu(x(t))/\mathrm{d}t^\nu = \partial_j\tilde{S}^\mu u^j_\nu = \delta^\mu_\nu ,$$

and therefore $|(\tilde{\Delta}^\mu_\nu)| = \tilde{s}_{(i)}v^{(i)} = 1$. Because on the other hand $v^{(i)}L_{(i)} = L = 1$, we must have $L_{(i)} = \tilde{s}_{(i)}$ on the extremals. As $\tilde{\Delta}^\mu_\nu = \delta^\mu_\nu$, the relations (7,28) become here

$$u^i_\mu = v^{(j)}\,\partial\tilde{s}_{(j)}/\partial(\partial_i S^\mu) , \tag{7,39a}$$

$$\partial L/\partial u^i_\mu = \partial_i\tilde{S}^\mu . \tag{7,39b}$$

Because of their properties the functions $\tilde{S}^\mu(x)$ can be used in order to solve the HJ eq. (7,24) off a given extremal, but such that on the extremal the eqs. (7,26) hold: Setting $S^{\bar{\mu}}(x) = \tilde{S}^{\bar{\mu}}(x)$, $\bar{\mu} = 2, \ldots, m$, we are left with a first order partial differential equation for $S(x) := S^1(x)$, which can be solved by solving its characteristic equations.

If $S(x; a)$ is a solution of the HJ eq. (7,24) and $x = f(t)$ an extremal for which eqs. (7,26) hold, then the current

$$G^\mu(t) = \frac{\partial S^\nu}{\partial a}\,\tilde{\Delta}^\mu_\nu(x = f(t)) \tag{7,40}$$

is conserved. The proof runs along the same lines as that for the current (5,109) in chapter 5 and will not be repeated here.

The Noether currents (5,114),

$$G^\mu = -T^\mu_\nu X^\nu + \pi^\mu_i Z^i$$

in our case have some special properties: Because of the relations (7,4a) the "energy-momentum" tensor $T^\mu_\nu = \pi^\mu_i u^i_\nu - \delta^\mu_\nu L$ vanishes identically for "parametric" systems. This does not mean that such systems cannot carry energy-momentum, but merely reflects the fact that the parameters t^μ are not physically relevant variables and that any change of them does not alter the physical (or "intrinsic" geometrical) properties of the system. Thus, the Noether currents here have the general form

$$G^\mu = \pi^\mu_i Z^i , \qquad Z^i = \left.\frac{\partial\hat{x}^i}{\partial a}\right|_{a=0} , \tag{7,41}$$

where $x^i \to \hat{x}^i(t, x; a)$, $\hat{x}^i(t, x; a = 0) = x^i$, $i = 1, \ldots, n$ is the symmetry transformation which leaves L invariant. The physical energy-momentum currents $T^\mu_i(t)$ are associated with the translations

$$x^i \to \hat{x}^i = x^i + a^i , \tag{7,42}$$

which, according to eq. (7,41), give the currents

$$T_i^\mu(t) = \pi_i^\mu(t)\,. \tag{7,43}$$

One sees immediately that the currents (7,43) are conserved only, $\partial_{(\mu)}\pi_i^\mu = 0$, if L does not depend on x, i.e. if $\partial_i L = 0$. In Minkowsky space $T_i^0 = \pi_i^0(t)$ is the energy current etc.

In many cases the DWHJ theory from chapter 4 can be used for "parametric" systems, too. The reason is that for $m \geq 2$ the relations (7,4b) in general do not imply that the Legendre transformation $u_\mu^i \to \pi_i^\mu$, $L \to H_{\mathrm{DW}}$ is singular, see, e.g., the relativistic string in chapter 4. Because of the eqs. (7.4a) we have

$$H_{\mathrm{DW}} = \pi_i^\mu u_\mu^i - L = (m-1)\,L\,. \tag{7,44}$$

The ansatz

$$S^\mu(t, x) = \frac{1-m}{m}\, t^\mu + \bar{S}^\mu(x) \tag{7,45}$$

reduces the DWHJ equation $\partial_{(\mu)}S^\mu + H_{\mathrm{DW}} = 0$ to $1 - m + H_{\mathrm{DW}}(x, \pi) = 0$, $\pi_i^\mu = \partial_i \bar{S}^\mu(x)$. Comparing eq. (7,44) with the HJ eq. (7,24) we see that

$$H_{\mathrm{DW}} = (m-1)\, H_{\mathrm{param.}} = m-1\,. \tag{7,46}$$

Furthermore, suppose the Lagrangian L does not depend on x, but only on the "velocities" u_μ^i, then, if $x = f(t)$ is an extremal, the functions

$$S^\mu(t, x) = \pi_i^\mu(t)\left(x^i - \frac{m-1}{m} f^i(t)\right) \tag{7,47}$$

are a solution of the DWHJ equation [Rinke, 1981]. The reason is that the Euler–Lagrange equations here have the form $\partial_{(\mu)}\pi_i^\mu(t) = 0$, so that

$$\partial_{(\mu)}S^\mu(t, x) = -\frac{m-1}{m}\, \pi_i^\mu u_\mu^i = -(m-1)\,L = -H_{\mathrm{DW}}.$$

7.4. Relativistic strings and electromagnetic fields of rank 2

There is an interesting relationship between the dynamics of relativistic strings $x^\alpha(t) = f^\alpha(t)$ in Minkowski space and Maxwell fields of rank 2 [Kastrup, 1978b, 1979; Rinke, 1980; Nambu, 1980; Kastrup and Rinke, 1981]: If the e.m. field 2-form $F = \frac{1}{2}F_{\alpha\beta}\, \mathrm{d}x^\alpha \wedge \mathrm{d}x^\beta$, $\mathrm{d}F = 0$, has rank 2, i.e. if $\varepsilon_{\alpha\beta\gamma\delta}F^{\alpha\beta}F^{\gamma\delta} = 0$, we have seen in chapter 1 that F has a representation $F = \mathrm{d}\bar{S}^1(x) \wedge \mathrm{d}\bar{S}^2(x)$, i.e. we have

$$F_{\alpha\beta}(x) = \partial_\alpha \bar{S}^1\, \partial_\beta \bar{S}^2 - \partial_\alpha \bar{S}^2\, \partial_\beta \bar{S}^1(x)\,. \tag{7,48}$$

The r.h. side of this equation has the same form as the functions $s_{\alpha\beta}(x)$ in eq. (7,25), so that we can identify $s_{\alpha\beta}(x)$ with $F_{\alpha\beta}(x)$, if $-\frac{1}{2}F^2 := -\frac{1}{2}F_{\alpha\beta}F^{\alpha\beta} = 1$. Suppose $F_{\alpha\beta}(x)$ is such a Maxwell field and suppose further that the functions $x^\alpha = f^\alpha(t)$ obey eq. (7,26),

$$v_{(\alpha\beta)}(t) = -L(v(t))\, F_{(\alpha\beta)}(x(t))\,, \tag{7,49}$$

then it follows from our discussion in section 7.3 that $x^\alpha = f^\alpha(t)$ is a solution of the field equations for the relativistic string. If $-\frac{1}{2}F^2 > 0$, but $-\frac{1}{2}F^2 \neq 1$, then the same conclusions hold for

$$s_{\alpha\beta}(x) = F_{\alpha\beta}(x)\,(-\tfrac{1}{2}F^2)^{-1/2}\,, \tag{7,50}$$

provided we have

$$\partial_\alpha[{}^*F^{\alpha\beta}(-\tfrac{1}{2}F^2)^{-1/2}] = 0\,, \qquad {}^*F_{\alpha\beta} = \tfrac{1}{2}\varepsilon_{\alpha\beta\gamma\delta}F^{\gamma\delta}\,. \tag{7,51}$$

The latter property is necessary, in order to guarantee that the form $s_{(\alpha\beta)}\, dx^\alpha \wedge dx^\beta$ is closed. Thus, given an e.m. field of rank 2 with the properties just mentioned, it determines a first order partial differential equation for relativistic strings!

Conversely, suppose $x^\alpha = f^\alpha(t)$ describes the motion of a relativistic string, with Plücker coordinates $v^{\alpha\beta}(t)$ and such that $L \neq 0$, then one can always find an electromagnetic field $F^{\alpha\beta}(x)$ of rank 2 with the property

$$F^{\alpha\beta}(x(t)) = \lambda\, v^{\alpha\beta}(t)\,, \qquad \lambda = \text{const.} \tag{7,52}$$

Proof: If the parametrization of the string is such that $L(v) = 1$, then it follows from the discussion in the last section 7.3 that it is always possible to find solutions $S^\mu(x)$, $\mu = 1, 2$, of the HJ eq. (7,25) such that for a given extremal $x = f(t)$ the relations

$$L_{(\alpha\beta)}(v(t)) = -v_{(\alpha\beta)}(t) = s_{(\alpha\beta)}(x)$$

hold! If we define

$$F^{\alpha\beta}(x) = -\lambda\, s^{\alpha\beta}(x)\,, \tag{7,53}$$

then $F_{\alpha\beta}$ has rank 2, because $\varepsilon_{\alpha\beta\gamma\delta}s^{\alpha\beta}s^{\gamma\delta} = 0$ and obeys the homogeneous Maxwell equations

$$\partial_\alpha {}^*F^{\alpha\beta}(x) = 0\,.$$

In general the functions $\partial_\alpha F^{\alpha\beta}$ will not vanish, but define the current $j^\beta(x)$. If $L(v) \neq 1, 0$, for the extremal $x = f(t)$, then one can proceed as follows: If one replaces the parameters t^μ by $\hat{t}^\mu = S^\mu(x(t))$, then $|\partial(\hat{t})/\partial(t)| = L(v) \neq 0$ and we can solve for $t^\mu = t^\mu(\hat{t})$. Inserting these functions into $L(v(t))$ gives $L(v) = \sigma(\hat{t})$. If we define $F_{\alpha\beta}(x)$ by

$$F_{\alpha\beta}(x) = -\lambda\sigma[S^1(x), S^2(x)]\, s_{\alpha\beta}(x)\,, \qquad -F^2 = \lambda^2\sigma^2(S^1, S^2)\,, \tag{7,54}$$

then not only the homogeneous Maxwell equations $\partial_\alpha {}^*F^{\alpha\beta} = 0$ are fulfilled, but eqs. (7,51) as well.

Example:
Consider the string motion

$$
\begin{aligned}
&x^0 = t^1\,, \qquad x^1 = A(t^2 - \tfrac{1}{2}\pi)\cos\omega t^1\,, \\
&x^2 = A(t^2 - \tfrac{1}{2}\pi)\sin\omega t^1\,, \qquad x^3 = 0\,, \\
&A\omega = 2/\pi\,, \qquad 0 \le t^2 \le \pi\,,
\end{aligned} \tag{7,55}
$$

with

$$
\begin{aligned}
&v^{01} = A\cos\omega t^1\,, \qquad v^{02} = A\sin\omega t^1\,, \qquad v^{03} = 0\,, \\
&v^{23} = 0\,, \qquad v^{31} = 0\,, \qquad v^{12} = -A(t^2 - \tfrac{1}{2}\pi)(2/\pi)
\end{aligned}
$$

and $L(v) = A(1-\omega^2\rho^2(t))^{1/2}$, $\rho(t) = [(x^1)^2 + (x^2)^2]^{1/2}(t)$. A solution of the HJ equation $[-s_{(\alpha\beta)}s^{(\alpha\beta)}]^{1/2} = 1$ with $v_{\alpha\beta}(t) = -L(v)\,s_{\alpha\beta}(x)$ is

$$
S^1(x) = (1-\omega^2\rho^2)^{-1/2}\,(x^0 - \omega\rho^2\,\mathrm{arctg}(x^2/x^1))\,, \qquad S^2(x) = \rho\,, \tag{7,56}
$$

from which we get, according to eqs. (7,54), the Maxwell fields

$$
\boldsymbol{E}(x) = -\lambda\,A(x^1, x^2, 0)/\rho\,, \qquad \boldsymbol{B}(x) = \lambda A(0,0,\omega\rho)\,. \tag{7,57}
$$

7.5. Bibliographical notes

Most of the relevant mathematical literature on systems which are invariant under reparametrization has already been quoted in the text above. "Parametric" variational problems are also discussed in the textbooks of Rado [1951], Funk [1970] and Rund [1973].

The interest of physicists in such systems came with the observation of Nambu [1970], Susskind [1970] and Nielsen [1970] that the relation between energy and angular momentum for the string is strikingly similar to the mass–spin dependence of mesons.

Possible relations between strings and electromagnetic fields have been discussed previously by Nielson and Oleson [1973], Kalb and Ramond [1974], Nambu [1974], Lund and Regge [1976], Englert and Windey [1978].

The simple relation (7,52) and the importance of rank 2 of the associated Maxwell field was suggested by myself [1978b, 1979] and systematically discussed in the framework of HJ theory by Rinke [1980] (my presentation above essentially follows Rinke's work). Shortly afterwards similar ideas were expressed by Nambu [1980] in a HJ framework which is essentially that of DeDonder–Weyl. See also Kastrup and Rinke [1981], Ogielski [1980], Nambu [1981] and Hosotani [1981].

8. The property $L = 0$ as a criterion for singularities in the transversality relations between HJ wave fronts and extremals

I mentioned already in chapter 5 that, despite its analytical complexities, Carathéodory's canonical theory for fields provides a powerful new handle for the analysis of *qualitative* properties of field theories. As a simple example we shall discuss in this chapter what happens, if the HJ wave fronts and the extremals "touch" each other, i.e. if the sum of their n- and m-dimensional tangent spaces at a point (x, z) no longer span an $(n+m)$-dimensional vector space. It was already mentioned at the end of chapter 3 that this breakdown of the general transversality properties happens if $H_c L = 0$. We shall mainly discuss here* the case $L = 0$ which provides us with an interesting criterion for bifurcations (phase transitions) in field theories [Kastrup, 1978a and b, 1981].

In order to get acquainted with the geometrical and physical meaning of the condition $L = 0$, we first investigate it in mechanics and then go on to field theories.

8.1. On the initial value problem for the HJ equation in mechanics

We have seen in chapter 2 that the determinant of the $(n+1)\times(n+1)$-matrix (2,39) formed by the tangent vector $e_t = \partial_t + \dot{q}^j\partial_j$ of an extremal and the n tangent vectors $w_{(j)} = p_j\partial_t + H\partial_j$ of a wave front at a point (t, q) has the value $-H^{n-1}L$. Therefore those $n+1$ tangent vectors are no longer linearly independent if $HL = 0$. Provided the wave fronts are indeed n-dimensional at this point, this "breakdown" of transversality means that the extremal is tangent to the wave front. In analogy to optics [Born and Wolf, 1975, p. 126; Guckenheimer, 1974] one might call such a point $(t_c, q_c) \in G^{n+1}$ a "focal" point and the set of all such points a "caustic":

Assume that the points of a light ray in R^3 are given by $\boldsymbol{x} = \boldsymbol{x}(s)$, $s \in [s_1, s_2]$ and the points of a 2-dimensional wave front by $\boldsymbol{x} = \boldsymbol{x}(u^1, u^2)$, $\boldsymbol{u} \in G^2$. Suppose the light ray $\boldsymbol{x}(s)$ transverses a given wave front at $\boldsymbol{x} = \boldsymbol{x}(\boldsymbol{u}_0)$, so that we can characterize that light ray by $\boldsymbol{x} = \boldsymbol{x}(s; \boldsymbol{u}_0)$. A light ray in the neighborhood of $\boldsymbol{x}(s; \boldsymbol{u}_0)$ may then be characterized by $\boldsymbol{u}_0 + \delta\boldsymbol{u}$ and one speaks of a focal point, if

$$\boldsymbol{x}(s+\delta s; \boldsymbol{u}_0 + \delta\boldsymbol{u}) = \boldsymbol{x}(s; \boldsymbol{u}_0)\,, \qquad \text{or} \qquad \frac{\partial \boldsymbol{x}}{\partial s}\,\delta s + \frac{\partial \boldsymbol{x}}{\partial u^1}\,\delta u^1 + \frac{\partial \boldsymbol{x}}{\partial u^2}\,\delta u^2 = 0\,,$$

which means that the tangent vectors $\partial\boldsymbol{x}/\partial s$ and $\partial\boldsymbol{x}/\partial u^\alpha$, $\alpha = 1, 2$ are linearly dependent at $(s, \boldsymbol{u}_0)$:

$$\det\left(\frac{\partial \boldsymbol{x}}{\partial s}, \frac{\partial \boldsymbol{x}}{\partial u^1}, \frac{\partial \boldsymbol{x}}{\partial u^2}\right)\bigg|_{\boldsymbol{u}=\boldsymbol{u}_0} = 0\,. \tag{8,1}$$

Before we discuss the physical meaning of the conditions $L = 0$ and $H = 0$, let me deal with the associated geometrical problem, following essentially Carathéodory [1937, ch. II]: Let $q^j = f^j(t; u^1, \ldots, u^\alpha, \ldots)$, $p_j(t; u^1, \ldots, u^\alpha, \ldots)$, $j = 1, \ldots, n$, be solutions of the canonical equations of motion, depending on the parameters u^α, where α takes at least the values $1, \ldots, n$, but can go up to $2n$. Then the function $\sigma(t; u)$ of eq. (2,12) obeys the differential equation

$$\partial_t\sigma = -H(t, f, g) + g_j\,\partial_t f^j\,. \tag{8,2}$$

Any solution $\sigma = s(t; u)$ is determined only up to an arbitrary function $A(u)$ of the parameters u. For the differential $\mathrm{d}s$ we obtain

*The case $H_c = 0$, or $|(T^\mu_\nu)| = 0$, will have to be investigated separately!

$$\begin{aligned} ds &= \partial_t s \, dt + \frac{\partial s}{\partial u^\alpha} du^\alpha = -H \, dt + g_j \, \partial_t f^j + \frac{\partial s}{\partial u^\alpha} du^\alpha \\ &= -H \, dt + g_j \, df^j + C_\alpha \, du^\alpha , \qquad C_\alpha = \partial s/\partial u^\alpha - g_j \, \partial f^j/\partial u^\alpha . \end{aligned} \tag{8,3}$$

Because of the relation (2,13) the coefficients C_α are independent of t and we have

$$\frac{\partial}{\partial u^\alpha} C_\beta - \frac{\partial}{\partial u^\beta} C_\alpha = [u^\alpha, u^\beta] . \tag{8,4}$$

As the solutions $\sigma = s(t; u)$ of eq. (8,2) are determined only up to a function $A(u)$, the coefficients $C_\alpha(u)$ have the corresponding arbitrariness. They become fixed, however, if we choose appropriate initial conditions: Suppose that for $t = \bar{t}(u)$ we have

$$\begin{aligned} &f^j(t = \bar{t}(u); u) = \bar{f}^j(u) , \qquad g_j(t = \bar{t}(u); u) = \bar{g}_j(u) , \\ &s(t = \bar{t}(u); u) = \bar{s}(u) , \end{aligned} \tag{8,5}$$

where the functions $\bar{f}^j(u)$, $\bar{g}_j(u)$ and $\bar{s}(u)$ are given. Then eq. (8,3) becomes for $t = \bar{t}$:

$$d\bar{s} - C_\alpha \, du^\alpha = -\bar{H} \, d\bar{t} + \bar{g}_j \, d\bar{f}^j , \qquad \bar{H} := H(\bar{t}, \bar{f}, \bar{g}) ,$$

and we obtain for $\sigma_0(t; u) = s(t; u) - \bar{s}(u)$

$$d\sigma_0 = -H \, dt + g_j \, df^j - (-\bar{H} \, d\bar{t} + \bar{g}_j \, d\bar{f}^j) . \tag{8,6}$$

The function $\sigma_0(t; u)$ is that solution of the differential eq. (8,2) which obeys the initial condition $\sigma_0(\bar{t}; u) = 0$. The coefficients C_α here have the form

$$C_\alpha = \bar{H} \, \partial\bar{t}/\partial u^\alpha - \bar{g}_j \, \partial\bar{f}^j/\partial u^\alpha . \tag{8,7}$$

The preceding considerations can be used in order to construct a solution $S(t, q)$ of the HJ equation $\partial_t S(t, q) + H(t, q, p = \partial S) = 0$, which on the n-dimensional surface $t = \bar{t}(u^1, \ldots, u^n)$, $q^j = \bar{f}^j(u^1, \ldots, u^n)$ has the given initial value

$$S(t = \bar{t}(u), q = \bar{f}(u)) = \bar{S}_0(u) . \tag{8,8}$$

Given the functions $\bar{t}(u)$, $\bar{f}^j(u)$ and $\bar{S}_0(u)$, the functions $\bar{g}_j(u)$ as defined in eqs. (8,5) are no longer arbitrary, because, according to chapter 2, the solutions of the canonical equations of motion generate a solution of the HJ equation, iff the Lagrange brackets $[u^j, u^k]$ vanish. According to eqs. (8,4) and (8,7) this is the case if

$$\partial\bar{S}_0/\partial u^j = -H(\bar{t}, \bar{f}, \bar{g}) \, \partial\bar{t}/\partial u^j + \bar{g}_k \, \partial\bar{f}^k/\partial u^j =: F_j(\bar{g}, u) . \tag{8,9}$$

The n functions $\bar{g}_j$ are to be determined from the n eqs. (8,9). These equations are solvable for $\bar{g}_j$, iff

$$\det\left(\frac{\partial F_j}{\partial \tilde{g}_k} = -\frac{\partial H}{\partial p_k}(\tilde{t}, \tilde{f}, \tilde{g})\frac{\partial \tilde{t}}{\partial u^j} + \frac{\partial \tilde{f}^k}{\partial u^j}\right) \neq 0 . \tag{8,10}$$

Suppose this is the case. Then we can determine the function $\sigma_0(t; u)$ discussed above and for the solution $\sigma(t; u) = \sigma_0(t; u) + \tilde{S}_0(u)$ of eq. (8,2) we have, according to eqs. (8,6) and (8,9),

$$d\sigma = -H\, dt + g_j\, df^j , \qquad \sigma(t = \tilde{t}, u) = \tilde{S}_0(u) . \tag{8,11}$$

Solving the equations $q^j = f^j(t; u)$ for the variables u^k, $k = 1, \ldots, n$, which is possible provided $\Delta(t; u) = |(\partial f^j/\partial u^k)| \neq 0$, we get $u^j = \chi^j(t, q)$. Inserting these functions into $\sigma(t; u)$ and $g_j(t; u)$ gives

$$S(t, q) = \sigma(t; u = \chi(t, q)) , \tag{8,12a}$$

$$\psi_j(t, q) = g_j(t; u = \chi(t, q)) \tag{8,12b}$$

and the relation (8,11) takes the form

$$dS(t, q) = -H[t, q, p = \psi(t, q)]\, dt + \psi_j(t, q)\, dq^j .$$

Thus the function $S(t, q)$ defined by (8,12a) obeys the HJ equation and has the initial value $S[t = \tilde{t}(u), q = \tilde{f}(u)] = \tilde{S}_0(u)$.

We here are mainly interested in the condition (8,10), which may be rewritten in the following way: Since

$$\frac{\partial \tilde{f}^k}{\partial u^j} = \left(\partial_t f^k \frac{\partial \tilde{t}}{\partial u^j} + \frac{\partial f^k}{\partial u^j}\right)_{t=\tilde{t}} = \left(\frac{\partial H}{\partial p^k}\frac{\partial \tilde{t}}{\partial u^j} + \frac{\partial f^k}{\partial u^j}\right)_{t=\tilde{t}}$$

the condition (8,10) is equivalent to

$$|(\partial f^k/\partial u^j)_{t=\tilde{t}}| \neq 0 . \tag{8,13}$$

The geometrical interpretation of the inequalities (8,10) or (8,13) is the following: The tangent vector of the extremal $(t, q = f(t; u))$ at $(t, q) = (\tilde{t}(u), \tilde{f}(u))$ is $\tilde{e}_t = (1, \partial_t f)_{t=\tilde{t}(u)}$ and the n-dimensional tangent space of the surface $t = \tilde{t}(u)$, $q^j = \tilde{f}^j(u)$ is spanned by the vectors $\tilde{w}_{(j)} = (\partial\tilde{t}/\partial u^j, \partial \tilde{f}^1/\partial u^j, \ldots)$. These $n+1$ vectors are linearly independent, if the determinant

$$|(\tilde{e}_t, \tilde{w}_{(1)}, \ldots, \tilde{w}_{(n)})| \neq 0 . \tag{8,14}$$

Again using formula (2,40b) and $\partial_t f^j = \partial H/\partial p_j$, we see that the determinant (8,14) is equal to the determinant (8,10).

Let us look at some special examples of the initial surfaces $t = \tilde{t}(u)$, $q^j = \tilde{f}^j(u)$.

(i) For $t = \tilde{t}(u) = t_0 = \text{const.}$ the eqs. (8,9) take the form

$$\tilde{g}_k\, \partial \tilde{f}^k/\partial u^j = \partial \tilde{S}_0/\partial u^j , \tag{8,15}$$

which can be solved for the functions $\tilde{g}_k$ iff

$$\left| \left(\frac{\partial \tilde{f}^k}{\partial u^j} = \frac{\partial f^k}{\partial u^j} \Big|_{t=t_0} \right) \right| \neq 0 . \tag{8,16}$$

The condition (8,16) is the conventional one that a family of extremals $q^j = f^j(t; u)$ does not have a focal point in the plane $t = t_0$ [Carathéodory, 1935, §327; Hestenes, 1966, ch. 3; DeWitt-Morette et al., 1979, App. B; Maslov and Fedoriuk, 1981].

If the inequality (8,16) holds, we can solve the equations $q^j = \tilde{f}^j(u)$ for $u^j = h^j(q)$ and obtain from eqs. (8,15):

$$\tilde{g}_j(u) = \frac{\partial \tilde{S}_0}{\partial u^k} \frac{\partial h^k}{\partial q^j} = \frac{\partial \tilde{S}_0}{\partial q^j} [u = h(q)] \Big|_{q=\tilde{f}(u)} . \tag{8,17}$$

These equations show that the initial surface $q^j = \tilde{f}^j(u)$, $p_j = \tilde{g}_j(u)$ forms a Lagrangian submanifold of the phase space (see section 2.2). The simplest choice for $\tilde{f}^j(u)$ obviously is $\tilde{f}^j(u) = u^j$ with $\tilde{g}_k = \partial \tilde{S}^0/\partial u^k$. If $u^j = \tilde{g}_j$, then eqs. (8,15) yield the consistency condition

$$\tilde{f}^j = \frac{\partial}{\partial u^j} \left(\sum_{k=1}^{n} u^k \tilde{f}^k - \tilde{S}_0(u) \right) .$$

(ii) Another possibility for choosing the function $\bar{t}(u)$ is

$$S_0(\bar{t}(u), u) = \tilde{S}_0(u) = \text{const.} \tag{8,18}$$

We now have

$$\frac{\partial \tilde{S}_0}{\partial u^j} = \frac{\partial S_0}{\partial \bar{t}} \frac{\partial \bar{t}}{\partial u^j} + \frac{\partial S_0}{\partial u^j} = 0 , \tag{8,19}$$

and eqs. (8,9) become

$$\tilde{g}_k \, \partial \tilde{f}^k / \partial u^j = \tilde{H} \, \partial \bar{t} / \partial u^j . \tag{8,20}$$

If we can solve the equations $q^j = \tilde{f}^j(u)$ for $u^j = h^j(q)$, then we obtain from eqs. (8,18) and (8,19)

$$\tilde{g}_k = -\tilde{H} \frac{\partial S_0}{\partial u^j} \frac{\partial h^j}{\partial q^k} \Big/ \frac{\partial S_0}{\partial \bar{t}} = \frac{\partial S_0}{\partial q^k} [\bar{t}, h(q)]_{q=\tilde{f}(u)} \tag{8,21}$$

because $\partial S_0/\partial \bar{t} = -\tilde{H}$ (relation (8,11) holds for the initial surface, too) which shows again that the initial surface is a Lagrangian submanifold of the phase space. Inserting the value (8,20) for $\partial \bar{t}/\partial u^j$ into the determinant (8,10) gives

$$\left|\left(-\partial_t f^k(t=\bar{t})\left(\bar{g}_l \frac{\partial \bar{f}^l}{\partial u^j}\Big/ \bar{H}\right)+\frac{\partial \bar{f}^k}{\partial u^j}\right)\right| = |(-\partial_t f^k(t=\bar{t})\,\bar{g}_l/\bar{H} + \delta^k_l)|\left|\left(\frac{\partial \bar{f}^k}{\partial u^j}\right)\right|$$
$$= -(\bar{L}/\bar{H})\,|(\partial \bar{f}^k/\partial u^j)|\,,\qquad \bar{L} = \partial_t f^k(t=\bar{t})\,\bar{g}_k - \bar{H}\,. \tag{8,22}$$

If we take $\bar{f}^j = u^j$, the determinant (8,22) is, up to a factor $\bar{H}^n$, the same as in eq. (2,41) if we take $t = \bar{t}$ there.

A simple example may serve to illustrate the above considerations: $L = (m/2)(\dot{x}^2 - \omega^2 x^2)$. The most general solution of the associated equation of motion is

$$x(t) = a\cos[\omega(t-\bar{t}) + \bar{\varphi}]\,,\qquad p(t) = m\,\dot{x}(t)\,,$$

with the initial values

$$x(\bar{t}) = \bar{f}(u) = a\cos\bar{\varphi}\,,\qquad p(\bar{t}) = \bar{g}(u) = -m\omega a\sin\bar{\varphi}\,.$$

From

$$\partial_t\sigma = L(x(t), \dot{x}(t)) = \tfrac{1}{2}m\dot{x}^2 - \tfrac{1}{2}m\omega^2 x^2$$
$$= -\tfrac{1}{2}m\omega^2 a^2\cos[2\omega(t-\bar{t}) + 2\bar{\varphi}]\,,$$

we get

$$\sigma(t;u) = \tfrac{1}{2}x(t;u)\,p(t;u) + A(u)\,, \tag{8,23}$$

where the constant of integration $A(u)$ depends on the choice of the initial data:

(i) If we choose the initial condition

$$t = \bar{t}(u) = t_0 = 0\,,\qquad x(t=\bar{t};u) = \bar{f}(u) = u\,,\qquad \bar{S}_0 = 0\,,$$

then we have

$$p(t=0;u) = -m\omega a\sin\bar{\varphi} = \bar{g}(u) = \partial\bar{S}_0/\partial u = 0\,,\qquad x(t=0;u) = a\cos\bar{\varphi}\,,$$

so that $\bar{\varphi} = 0$, $x(t) = a\cos\omega t$ and

$$\sigma(t;u) = -\tfrac{1}{2}m\omega a^2\cos\omega t\sin\omega t\,. \tag{8,24}$$

Eliminating $a = x/\cos\omega t$ gives

$$S(t,x) = -\tfrac{1}{2}m\omega x^2\,\mathrm{tg}\,\omega t\,, \tag{8,25}$$

which has

$$S(t=0,x) = 0\,,\qquad p = \partial_x S = -m\omega x\,\mathrm{tg}\,\omega t\,,$$

so that indeed

$$\frac{1}{2m}[(\partial_x S)^2 + m\omega^2 x^2] = -\partial_t S .$$

The family of solutions $x = f(t; a) = a \cos \omega t$, parametrized by a, has a "focal" point at $\omega t = \omega t_c = \pi/2$, because $(\mathrm{d}x/\mathrm{d}a)(t = t_c) = 0$. On the other hand, the Lagrangian for the solution $x = a \cos \omega t$ is

$$L = \tfrac{1}{2}m\omega^2 a^2 (1 - 2\cos^2 \omega t) .$$

It vanishes at the point $(\omega \bar{t} = \pi/4,\ \bar{x} = a/\sqrt{2})$. A wave front passing through this point is

$$S(t, x) = -\tfrac{1}{2}m\omega x^2 \operatorname{tg} \omega t = -\tfrac{1}{4}m\omega a^2 .$$

A tangent vector of the extremal $(t, x(t))$ at the point $\omega \bar{t} = \pi/4$ is $(1, -a\omega/\sqrt{2})$ and a tangent vector of the wave front is $w = (-\partial_x S/\partial_t S,\ 1) = (-\sqrt{2}/(a\omega),\ 1)$, so that extremal and wave front indeed touch each other at the point $(\omega \bar{t} = \pi/4,\ \bar{x} = a/\sqrt{2})$. However, this point does not show any interesting intrinsic dynamical significance, contrary to the point $(\omega t,\ x) = (\pi/2,\ 0)$. One reason for this is that the normalization of the potential energy (or of the total energy) is arbitrary: For instance, if we replace the potential $\frac{1}{2}m\omega^2 x^2$ by $\frac{1}{2}m\omega^2 x^2 + \frac{1}{2}m\omega^2 a^2$, the equations of motion and the initial condition do not change. However, now the Lagrangian has the form $L = -m\omega^2 a^2 \cos^2 \omega t$ for the solution $x = a \cos \omega t$ and this L vanishes at $\omega t_c = \pi/2$. We shall come back to this normalization problem in the next paragraph.

(ii) Suppose the initial data are

$$\begin{aligned} &x(t = \bar{t}; u) = a \cos \bar{\varphi} = \bar{f}(u) = u , \\ &S_0(\bar{t}(u), u) = -\tfrac{1}{2}m\omega^2 a^2 (\bar{t} - \tau(u)) = \bar{S}_0 = 0 , \end{aligned} \qquad (8{,}26)$$

then we have, according to eq. (8,21),

$$p(t = \bar{t}, u) = -m\omega a \sin \bar{\varphi} = \bar{g}(u) = \left.\frac{\partial S_0}{\partial u}\right|_{S_0 = 0} = \tfrac{1}{2}m\omega^2 a^2 \tau' , \qquad \tau' = \frac{\mathrm{d}\tau}{\mathrm{d}u} ,$$

so that

$$a^2(u) = \frac{2}{\tau'^2\omega^2}[1 \pm (1 - \omega^2\tau'^2 u^2)^{1/2}] , \qquad \sin 2\bar{\varphi} = -\omega\tau' u . \qquad (8{,}27)$$

For the function $\sigma(t; u)$ we now get

$$\sigma(t; u) = \tfrac{1}{2}x(t)\, p(t) - \tfrac{1}{2}u\, \bar{g}(u) - \tfrac{1}{2}m\omega^2 a^2 (\bar{t} - \tau(u)) . \qquad (8{,}28)$$

Eliminating $\bar{t}$ on the r.h. side of eq. (8,28) with the help of $x = a \cos[\omega(t - \bar{t}) + \bar{\varphi}]$ gives

$$S(t, x) = \tfrac{1}{2}m\omega x(a^2 - x^2)^{1/2} - \tfrac{1}{2}m\omega^2 a^2\left(t - \frac{1}{\omega}\arccos\frac{x}{a}\right) - \tfrac{1}{2}u\, \bar{g}(u) + \tfrac{1}{2}m\omega^2 a^2\left[\tau(u) - \frac{1}{\omega}\bar{\varphi}(u)\right] . \qquad (8{,}29)$$

Because it satisfies

$$\partial_t S = -\tfrac{1}{2}m\omega^2 a^2 = -E\,, \qquad \partial S/\partial x = -m\omega\,(a^2 - x^2)^{1/2}\,,$$

the function (8,29) is a solution of the HJ equation with the initial value $S(t = \tilde{t}(u),\ x = u) = 0$.

The focal points of the solution $x = f(t; u) = a(u)\cos[\omega(t - \tilde{t}(u)) + \tilde{\varphi}(u)]$ here, too, are given by the zeros of $\mathrm{d}f/\mathrm{d}u$. They will in general not coincide with the points $(t, \bar{q})$ for which $L = 0$. Thus, it appears necessary to clarify the intrinsic meaning of the condition "$L = 0$".

8.2. *The physical meaning of the property $L = 0$ in mechanics*

We have seen above that the HJ wave fronts and particle trajectories become tangent, if either $H = 0$ or $L = 0$. In order to analyze the question, when these conditions have an intrinsic dynamical meaning, we look at the Lagrangian

$$L = \tfrac{1}{2}m(\dot{r}^2 + r^2\dot{\varphi}^2) - V(r) \tag{8,30}$$

of a particle in R^3, moving in a rotationally symmetric potential. We first take the normalization of the potential $V(r)$ to be the conventional one: $V(r = 0) = 0$ if $V(r)$ is finite at the origin and $V(r \to \infty) = 0$, if $V(r)$ is singular at $r = 0$ and finite at infinity. Then the condition $H = E = 0$ separates the bounded from the unbounded motions (if both are possible in principle), i.e. the condition $E = 0$ defines a bifurcation line in the (l, E)-plane of initial values, parametrized by the angular momentum l and the energy E [Pars, 1965, ch. 17; Smale, 1970]. This obvious interpretation of the condition $H = 0$ leads us to look at the condition "$L = 0$" from the following point of view [Kastrup, 1981]: If $q^j = f^j(t, u)$, $j = 1, \ldots, n$, is a solution of the equations of motion, then

$$\Lambda(t; u) = L[q = f(t; u),\ \dot{q} = \partial_t f(t; u)]$$

may vanish for certain $t = \bar{t}$. We have seen above that such points in general do not have an intrinsic physical meaning. However, it may happen that $\Lambda(t; u) = 0$ *for all t*, if the set $(u^1, \ldots)$ of the parameters u which determine the initial conditions takes a certain value. We can give an intrinsic meaning to this condition in the following way:

Assume that for a certain choice of the parameters $(u^1, \ldots)$ we have $\Lambda(t; u) = \lambda = \text{const.}$, i.e. the Lagrangian is a constant of the particular motion under consideration. If $H = E = \text{const.}$, then we can "renormalize" the energy $E \to E + \lambda$ such that $\Lambda \to \Lambda - \lambda = 0$. Thus, the relevant interpretation of the condition "$L = 0$" for all t, which avoids any normalization ambiguities of the energy, is

$$\frac{\mathrm{d}}{\mathrm{d}t}L = 0\,, \qquad \text{or} \qquad \frac{\mathrm{d}}{\mathrm{d}t}(p_j\dot{q}^j) = 0\,. \tag{8,31}$$

If $L = T - V$ this means that kinetic energy T and potential energy V are constants of motion separately! This is possible if the particles move along equipotential lines.

For the Lagrangian (8,30) the condition $L = \text{const.}$ defines a bifurcation set of points $(l, E_0(l))$ in the (l, E)-plane, too: $V(r) = \text{const.}$ means that $r = r_0 = \text{const.}$ and this, in general, is just the motion at a (local) minimum of the effective potential $V_{\text{eff}}(r) = l^2/(2mr^2) + V(r)$ (the motion at a local maximum is unstable, i.e. r_0 is a solution of the equation $\mathrm{d}\,V_{\mathrm{eff}}/\mathrm{d}r\ \ (r=r_0)=0$:

$$r_0^3\,V'(r_0)=l^2/m\,. \tag{8,32}$$

Depending on the number of critical points of $V_{\mathrm{eff}}(r)$, there can be several solutions $r_0^{(i)}$, $i=1,\ldots,$ of eq. (8,32). Notice that the value of r_0 is independent of the normalization of the potential $V(r)$ – or the energy E –, as it should. The roots r_0 will be functions of the angular momentum l and the parameters of the potential $V(r)$. For the power potential $V(r)=gr^{\nu}$, $\nu>-2$, we get

$$r_0=(l^2/\nu g m)^{1/(\nu+2)}=r_0(l)\,, \tag{8,33}$$

where $\nu g>0$ is assumed. For $r=r_0$ we have $\dot r_0=0$, so that the total energy E_0 becomes

$$E_0=E_0(l)=l^2/(2mr_0^2)+V(r_0)\,. \tag{8,34}$$

For the potential $V(r)=gr^{\nu}$, for instance, we get

$$E_0(l)=\frac{\nu+2}{2\nu}(\nu g)^{2/(\nu+2)}\left(\frac{l^2}{m}\right)^{\nu/(\nu+2)}. \tag{8,35}$$

The curve $E_0=E_0(l)$ in the (l,E)-plane again defines a bifurcation set: If r_0 describes a local minimum of $V_{\mathrm{eff}}(r)$, then, for energies E "slightly above" E_0, $E-E_0=\varepsilon>0$ small, we will have a bounded motion $r(t)$ with two turning points, $r_+\geq r(t)\geq r_-$, $r_+>r_0>r_-$, but for $E<E_0$ there is no motion at all! (If r_0 corresponds to a local maximum of $V_{\mathrm{eff}}(r)$, the motion for $r<r_0$, $E<E_0$ will be bounded, that for $r>r_0$, $E<E_0$, or $E>E_0$, will be bounded or unbounded, depending on the shape of the potential.)

In any case we see that the curves in the (l,E)-plane defined by $\mathrm{d}\Delta(t;E,l)/\mathrm{d}t=0$ separate qualitatively different "phases" of all possible motions associated with the potential $V(r)$.

At this point let me mention – very briefly – the main idea of the modern theories of turning points and caustics, in the framework of HJ theories and, more generally, Lagrangian submanifolds [Keller, 1958; Arnold, 1967 and 1978, appendices 11 and 12; Maslov, 1972; Berry and Mount, 1972; Duistermaat, 1974; Maslov and Fedoriuk, 1981; Berry, 1981]: In HJ theories turning points are characterized by $p_j=\partial_j S(t,q)=0$. Thus, they are critical points of the function $S(t,q)_{q=\mathrm{const.}}$, which locally defines an n-dimensional Lagrangian manifold $L_t^{(n)}$ with local coordinates $(q,\ p=\partial S(t,q))$. Thus, turning points correspond to singularities of the map which projects the Lagrangian manifold $L_t^{(n)}$ on the configuration space $\{q\}$. The set of such singular points is called a "caustic".

Since the radius r_0 of the circular motion considered above can be looked at as the limit of a motion the two turning points r_- and r_+ of which coalesce, our condition $L=\mathrm{const.}$ appears to be related to the "established" theory of caustics. The limit $r\to r_0$, or $E\to E_0(l)$ is, however, not without delicacies: Take, for instance, the determinant

$$\Delta(t;l,E)=\frac{\partial r}{\partial l}\frac{\partial\varphi}{\partial E}-\frac{\partial r}{\partial E}\frac{\partial\varphi}{\partial l}\,,$$

the zeros of which give the focal points of the motion. If we insert $E_0=E_0(l)$ into $r=r(t;l,E)$ and $\varphi(t;l,E)$ before calculating the derivatives $\partial r/\partial l$ etc., then $\Delta(t;l,E_0(l))=0$ for all t, because the functions

$r(t; l) = r(t; l, E_0(l))$, $\varphi(t; l) = \varphi(t; l, E_0(l))$ are clearly dependent. If, on the other hand, we calculate $\Delta(t; l, E)$ first, and then the limit $E \to E_0(l)$, we encounter singularities: Differentiating

$$\frac{m}{2}\dot{r}^2 + \frac{l^2}{2mr^2} + V(r) = E$$

with respect to E gives

$$m\dot{r}\frac{\partial \dot{r}}{\partial E} + \left(-\frac{l^2}{mr^3} + V'(r)\right)\frac{\partial r}{\partial E} = 1 . \tag{8,36}$$

Because $\dot{r} \to 0$, $(-l^2/(mr^3) + V'(r)) \to 0$ for $r \to r_0$, then, according to eq. (8,36), the derivatives $\partial\dot{r}/\partial E$, $\partial r/\partial E$ cannot remain finite in this limit.
Example: Harmonic oscillator in the plane, $V = \frac{1}{2}m\omega^2 r^2$, with the solution

$$x = a\cos\omega t , \qquad y = b\sin\omega t ,$$

$$a(l, E) = (m\omega^2)^{-1/2}[E + (E^2 - \omega^2 l^2)^{1/2}]^{1/2} ,$$

$$b(l, E) = (m\omega^2)^{-1/2}[E - (E^2 - \omega^2 l^2)^{1/2}]^{1/2} .$$

The limit $r \to r_0$ is equivalent to $E \to \omega l$. One sees that the derivatives $\partial a/\partial E$, $\partial a/\partial l$ etc. become singular in this limit. $\Delta(t; l, E)$ becomes singular, too.

The functions $E_0(l)$ of eqs. (8,34) and (8,35) represent the ground state energies of the system for a given angular momentum. The "bifurcation" functions $E_0(l)$ have a remarkable resemblance to the corresponding quantum mechanical ground state energy levels with vanishing radial quantum number $n_r = 0$. Take, for instance, the harmonic oscillator ($\nu = 2$) and the Coulomb potential ($\nu = -1$). Here the functions (8,35) become

$$E_0(l) = \omega l , \qquad \omega = (2g/m)^{1/2} , \tag{8,37a}$$

$$E_0(l) = -\tfrac{1}{2}mg^2 l^{-2} , \tag{8,37b}$$

which differ from the corresponding quantum mechanical values

$$E_0^{\mathrm{qm}}(l) = \omega\hbar(l + \tfrac{3}{2}) , \qquad l = 0, 1, \ldots \tag{8,38a}$$

$$E_0^{\mathrm{qm}}(l) = -\tfrac{1}{2}mg^2[\hbar(l+1)]^{-2} , \qquad l = 0, 1, \ldots \tag{8,38b}$$

essentially by the zero-point energies which can be taken into account by replacing the classical value l by $(l + \lambda)$, where $\lambda = \frac{3}{2}$ for the harmonic oscillator and $\lambda = 1$ for the Coulomb potential. More generally: for $\nu > 0$ the WKB-approximations $E_0^{\mathrm{WKB}}(l)$ for the $n_r = 0$ energy levels are given by Quigg and Rosner [1979, ch. 4]:

$$\begin{aligned} E_0^{\mathrm{WKB}}(l) &= [C(\nu)]^{2\nu/(2+\nu)} g^{2/(2+\nu)} [\hbar^2(l + \tfrac{3}{2})^2/m]^{\nu/(2+\nu)} , \\ C(\nu) &= \nu\sqrt{\pi/2}\,\Gamma(3/2 + 1/\nu)/\Gamma(1/\nu) , \end{aligned} \tag{8,39}$$

which again have the same g-dependence as the bifurcation functions (8,35) and the same l-dependence, if we replace l by $\pi(l+\frac{3}{2})$. If we make this replacement and denote the resulting functions by $\tilde{E}_0(l)$, then these functions differ from the WKB-expressions (8,39) merely by a normalization factor which depends only on ν. Thus, for fixed ν, the ratios $\tilde{E}_0(l_1)/\tilde{E}_0(l_2)$ and $E_0^{\mathrm{WKB}}(l_1)/E_0^{\mathrm{WKB}}(l_2)$ are the same.

Another amusing example is provided by the logarithmic potential $V(r) = g\ \ln(r/a)$, for which we get from the eqs. (8,32) and (8,34) $r_0 = l(mg)^{-1/2}$ and

$$E_0(l) = g(\tfrac{1}{2} + \ln[l/(a\sqrt{mg})])\,. \tag{8,40}$$

Notice that the difference $E_0(l_2) - E_0(l_1) = g\ \ln(l_2/l_1)$ does not depend on a and m. The expression (8,40) can be compared with the WKB-approximation suggested by Quigg and Rosner [1979, p. 206]:

$$E_0^{\mathrm{WKB}}(l) = g(\tfrac{1}{2}\ln(\pi/2) + \ln[\hbar(l+\tfrac{3}{2})/(a\sqrt{mg})])\,. \tag{8,41}$$

Thus, if we replace l in eq. (8,40) by $\hbar(l+\frac{3}{2})$, the resulting energy levels $E_0(l)$ are higher than those given by eq. (8,41), because $\ln(\pi/2) = 0.451\ldots$. Whereas formula (8,41) gives values which are below the real ones [Quigg and Rosner, 1979, table 6 on p. 206], $\tilde{E}_0(l) = g(\frac{1}{2} + \ln[\hbar(l+\frac{3}{2})/(a\sqrt{mg})])$ gives energy levels above the real ones.

These examples strongly suggest that the bifurcation functions $E_0(l)$ of the classical systems provide – after a suitable correction which takes care of the zero-point energy – reasonable approximations for the corresponding quantum mechanical energy levels with vanishing radial quantum number. This conjecture is supported by comparing the analytical expression $E_0(l)$ for the anharmonic potential $V(r) = \frac{1}{2}b_1r^2 + \frac{1}{4}b_2r^4$ with the corresponding quantum mechanical levels [Kastrup, 1981].

Two other interesting examples with $L = \text{const.}$ are the following: For a particle moving in an external magnetic field $\boldsymbol{B} = \text{curl}\,\boldsymbol{A}$ the Lagrangian is $L = (m/2)\,\boldsymbol{v}^2 + q\,\boldsymbol{v}\cdot\boldsymbol{A}$. As $\boldsymbol{v}^2 = \text{const.}$ for a motion in a static external magnetic field, we have $L = \text{const.}$, if $\boldsymbol{v} = \kappa\boldsymbol{A}$, $\kappa = \text{const.}$ This happens in the following cases:

(i) $\boldsymbol{B} = \textbf{const.}$, $\boldsymbol{A} = \frac{1}{2}\boldsymbol{B}\wedge\boldsymbol{x}$. It follows from

$$m\ \mathrm{d}\boldsymbol{v}/\mathrm{d}t = q\boldsymbol{v}\wedge\boldsymbol{B} = q\frac{\mathrm{d}}{\mathrm{d}t}(\boldsymbol{x}\wedge\boldsymbol{B})$$

that $m\boldsymbol{v} = q\boldsymbol{x}\wedge\boldsymbol{B} = -2q\boldsymbol{A}$, assuming that $\boldsymbol{v} = 0$ for $\boldsymbol{x}$ parallel to $\boldsymbol{B}$. Notice that $L = 0$ in this example.

(ii) The Meissner effect ("Higgs mechanism") in a superconductor can be explained by the assumption [e.g. Kittel, 1971, p. 424] that the canonical momentum vanishes: $\boldsymbol{p} = m\boldsymbol{v} + q\boldsymbol{A} = 0$, i.e. we have $\boldsymbol{v} = -(q/m)\,\boldsymbol{A}$ in this case.

We observe that for $\boldsymbol{v} = \kappa\boldsymbol{A}$ the circulation $Z_c = \oint \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{x}$ along a closed curve C is given by

$$Z_c = \kappa\int_{F_c}\boldsymbol{B}\cdot\mathrm{d}\boldsymbol{f} = \kappa\Phi_c\,,$$

where Φ_c is the magnetic flux through a surface F_c which has the curve C as its boundary.

8.3. The property $LH_c = 0$ as a criterion for transversality singularities in field theories

The condition $HL \neq 0$ for the linear independence of the tangent vectors of extremals and wave fronts in mechanical systems takes the same form in Carathéodory's canonical theory, as can be seen as follows: According to section 5.4 the tangent space of the wave front $S^\mu(x, z) = \sigma^\mu = \text{const.}$, $\mu = 1, \dots, m$, at $(x, z) \in G^{m+n}$ can be spanned by the n vectors

$$w_{(a)} = H_c(k^\mu_a \partial_\mu + \partial_a) = p^\mu_a \partial_\mu + H_c \partial_a\,, \tag{8,42}$$

and the tangent space of an extremal $z^a = f^a(x)$ can be spanned by

$$e_{(\mu)} = \partial_\mu + v^a_\mu \partial_a\,, \qquad v^a_\mu = \partial_\mu f^a(x)\,. \tag{8,43}$$

The $n + m$ vectors (8,42) and (8,43) are linearly independent, if the determinant

$$\left|\begin{pmatrix} E_m & p \\ v & H_c E_n \end{pmatrix}\right|\,, \qquad p = (p^\mu_a)\,, \qquad v = (v^a_\mu)\,,$$

does not vanish. According to formula (2,40b) this determinant equals

$$|(H_c E_n - v \cdot p)| = (-1)^n\,|Q|$$

and therefore, because of eq. (5,22), we have

$$\left|\begin{pmatrix} E_m & p \\ v & H_c E_n \end{pmatrix}\right| = -H_c^{n-1} L\,. \tag{8,44}$$

Thus, again as in mechanics, the transversality relations between extremals and CHJ wave fronts become singular if $H_c = 0$ or $L = 0$!

In order to get a feeling for the physical relevance of the condition $H_c L = 0$, we best look at a number of examples:

We first observe that – contrary to mechanics, where $H = \text{const.}$ if L does not depend explicitly on t – the Hamilton function H_c in field theories in general will not be a constant for solutions $z^a = f^a(x)$ of the field equations, even if L does not depend explicitly on x. This can already be seen in the case of one real scalar field in Minkowski space where $H_c = \pi^\mu v_\mu - L$. Because $\mathrm{d}\pi^\mu/\mathrm{d}x^\mu = \partial L/\partial z$ we have

$$\frac{\mathrm{d}}{\mathrm{d}x^\nu} L = \frac{\partial L}{\partial z}\,\partial_\nu f + \pi^\mu \partial_\nu \partial_\mu f = \frac{\mathrm{d}}{\mathrm{d}x^\mu}(\pi^\mu \partial_\nu f)\,,$$

and therefore

$$\frac{\mathrm{d}}{\mathrm{d}x^\nu} H_c = \left(\frac{\mathrm{d}}{\mathrm{d}x^\nu}\,\pi^\mu\right)\partial_\mu f - \left(\frac{\mathrm{d}\pi^\mu}{\mathrm{d}x^\mu}\right)\partial_\nu f\,, \tag{8,45}$$

which in general will not vanish. It does vanish for the special case that $L = \frac{1}{2} v_\mu v^\mu - V(z)$, $\pi^\mu = v^\mu$ and solutions of the type $z = f(x) = f(k \cdot x)$, $k = (k^\mu) = \text{const.}$, exist.

The Lagrangian (5,55) for the string is constant on the extremals [Schild, 1977] and since $H_c = -L$ for that system, H_c is constant, too.

8.4. Examples of systems in Minkowski space which have solutions of the field equations with $L = 0$, especially gauge theories with $\boldsymbol{E}^2 = \boldsymbol{B}^2$

There are important Lagrangians in physics which vanish for any solution of the field equations: The Lagrangian $L = \sqrt{-g}\, R$ of General Relativity [Landau and Lifshitz, vol. 2, ch. 11] vanishes where the trace θ^μ_μ of the energy-momentum tensor vanishes, for instance in regions, where there is no matter at all.

Another example of this kind is the Lagrangian (6.13) for the free Dirac equation which vanishes for every solution of the field equation

$$\mathrm{i}\gamma^\mu \partial_\mu \psi - m\psi = 0\,, \qquad \mathrm{i}\partial_\mu \bar{\psi}\gamma^\mu + m\bar{\psi} = 0\,.$$

More instructive is the case, when the fermion fields $\psi(x)$ are coupled to a gauge potential $A_\mu = A^a_\mu t_a$, where t_a is an element of the Lie algebra of the group $\mathrm{SU}(n)$. The associated gauge fields are

$$F_{\mu\nu} = F^a_{\mu\nu} t_a = \partial_\mu A_\nu - \partial_\nu A_\mu + \mathrm{i}g[A_\mu, A_\nu]\,.$$

If T_a is a representation of the elements t_a of the Lie algebra which acts on the spinor fields $\psi(x)$, then the Lagrangian of the coupled system is

$$\begin{aligned} L &= -\tfrac{1}{4}\operatorname{tr}(F_{\mu\nu}F^{\mu\nu}) + \tfrac{1}{2}\mathrm{i}(\bar{\psi}\gamma^\mu \mathrm{D}_\mu\psi - \overline{\mathrm{D}_\mu\psi}\gamma^\mu\psi) - m\bar{\psi}\psi\,, \\ \mathrm{D}_\mu &:= \partial_\mu - \mathrm{i}gA^a_\mu T_a\,. \end{aligned} \tag{8,46}$$

For solutions of the field equations the Lagrangian L becomes

$$L = L_{\text{extr.}} = -\tfrac{1}{4}\operatorname{tr}(F_{\mu\nu}F^{\mu\nu}) = \frac{1}{2}\sum_{a=1}^{n} [(\boldsymbol{E}^a)^2 - (\boldsymbol{B}^a)^2] =: \tfrac{1}{2}(\boldsymbol{E}^2 - \boldsymbol{B}^2)\,. \tag{8,47}$$

Thus we have $L = 0$ whenever the electric and magnetic energy densities are equal! Solutions with this property constitute a bifurcation hypersurface in the space of field configurations which separates electric and magnetic "phases" of the system. Notice that $L = 0$ is a Lorentz-invariant statement.

In E-dynamics the radiation field has the property $\boldsymbol{E}^2 - \boldsymbol{B}^2 = 0$. In nonabelian gauge theories Coleman's generalized plane waves [Coleman, 1977] have this property, too. In addition, any time-dependent (classical) solution with localizable finite energy of a pure nonabelian gauge theory has the property $\lim_{t\to\infty}(\boldsymbol{E}^2_a - \boldsymbol{B}^2_a) = 0$ for each internal index a [Glassey and Strauss, 1980; Magg, 1982]. See also a recent paper by Kyriakopoulos [1982].

An intriguing question is, whether there are nontrivial (i.e. genuinely nonabelian) solutions of the Yang–Mills equation with the property $\boldsymbol{E}^2 = \boldsymbol{B}^2$. This is a first order condition (for the potentials), which may be rewritten in the following way: If we define the dual fields by ${}^*F_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\rho\sigma}F^{\rho\sigma}$, then the condition $\boldsymbol{E}^2 = \boldsymbol{B}^2$ is equivalent to ${}^*\boldsymbol{E}^2 = \boldsymbol{E}^2$ or ${}^*\boldsymbol{B}^2 = \boldsymbol{B}^2$. This Minkowski space condition is strongly reminiscent of the condition ${}^*F^{\mu\nu}_E = \pm F^{\mu\nu}_E$ characteristic for "instanton" solutions in euclidean nonabelian gauge theories [Belavin et al., 1975; Coleman, 1979; Eguchi et al., 1980 (review with many references)]. In the euclidean case we have $L_E = -\frac{1}{4}F^{\mu\nu}_E F^{\mu\nu}_E = -\frac{1}{2}(\boldsymbol{E}^2_E + \boldsymbol{B}^2_E)$ and the energy density

$\theta^{00} = \frac{1}{2}(\boldsymbol{B}_E^2 - \boldsymbol{E}_E^2)$. The passage from the euclidean to Minkowski space is done by replacing x^4 by $\mathrm{i}x^0$, A_E^4 by $\mathrm{i}A^0$ and F_E^{k4} by $\mathrm{i}F^{k0}$, so that $\boldsymbol{B}_E^2 \to \boldsymbol{B}^2$, $\boldsymbol{E}_E^2 \to -\boldsymbol{E}^2$.

Thus, there is a striking similarity between the condition $\boldsymbol{E}^2 = \boldsymbol{B}^2$ ($L = 0$) in Minkowski space and the instanton property $\boldsymbol{E}_E^2 = \boldsymbol{B}_E^2$ in euclidean space. In physical applications instantons $A_\mu^{(\mathrm{inst.})}$ serve as external "background" fields for Dirac spinors $\psi(x)$ obeying the equations $\gamma^\mu(\partial_\mu - \mathrm{i}gA_\mu^{(\mathrm{inst.})})\,\psi = 0$. The solutions $\psi(x)$ of these equations appear to be of considerable interest for the analysis of chiral symmetry breaking ['t Hooft, 1976a; Coleman, 1979; Callan Jr. et al., 1979; Jackiw, 1980] and for the calculation of small quantum fluctuations around the classical instanton field ['t Hooft, 1976b; Brown et al., 1977; Bashilov and Pokrovsky, 1978; Schwarz, 1979 (with many refs.)].

Considering the interesting physical applications of the (euclidean) instanton solutions, the question arises whether classical genuinely "nonabelian" solutions with the property $\boldsymbol{E}^2 = \boldsymbol{B}^2$ of Yang–Mills theories in Minkowski space would not be similarly useful. Unfortunately, no such solutions are known – at least not to me! However, already the abelian case indicates that such solutions may be of considerable value: Let $A^\mu = A^\mu(\varphi)$, $\varphi = k_\mu x^\mu$, $k^2 = 0$, be the vector potential for a radiation field in the Lorentz gauge, $\partial_\mu A^\mu = 0$, or $k_\mu A^\mu = 0$. For such a field we have $\boldsymbol{E}^2 = \boldsymbol{B}^2$. If we put Dirac particles with charge e in such an external field, Dirac's equation

$$\gamma^\mu(\mathrm{i}\partial_\mu - eA_\mu)\,\psi - m\psi = 0$$

can be solved exactly [Wolkow, 1935; Berestetskij et al., 1971, §40]:

$$\psi(x) = \left(1 + \frac{e}{2k\cdot p}\not{k}\not{A}\right)u(p)\,\mathrm{e}^{\mathrm{i}S}\,, \qquad \not{a} := a_\mu\gamma^\mu\,,$$
$$S = -p\cdot x - \frac{1}{k\cdot p}\int_0^{k\cdot x}(e\,p\cdot A - \tfrac{1}{2}e^2A^2)\,\mathrm{d}\varphi\,, \tag{8,48}$$

where $p^2 = m^2$ and $(\not{p} - m)\,u = 0$. Of special interest here is the expression for the current $e\bar{\psi}\gamma^\mu\psi$ resulting from solution (8,48). It has the form

$$j^\mu(x) = \frac{e}{m}\,\bar{u}\cdot u[p^\mu - eA^\mu + k^\mu(e\,p\cdot A - \tfrac{1}{2}e^2A^2)/k\cdot p]\,. \tag{8,49}$$

The second term $j_A^\mu = -(e^2/m)\bar{u}uA^\mu$ on the r.h. side of eq. (8,49) is remarkable, because it may be interpreted as an effective photon mass term (with mass $\mu = (e^2\bar{u}u/m)^{1/2}$) induced by the interaction with the Dirac particles. This interpretation is not so strange as it looks at first sight:

According to Schwinger [1962, 1963] an external vector potential A^μ induces in 2-dimensional QED a current

$$j^\mu = -\frac{e^2}{\pi}A^\mu + \partial^\mu\left(\frac{e^2}{4\pi}\operatorname{tr}\varphi\right),$$
$$\Box\operatorname{tr}\varphi = 4\partial_\mu A^\mu\,. \tag{8,50}$$

The first term $j_A^\mu = -(e^2/\pi)\,A^\mu$ indeed represents a dynamically induced mass $(e^2/\pi)^{1/2}$ of the vector field A^μ. Anderson [1963] pointed out that Schwinger's consideration may be applied to the interaction of photons with an electron plasma in a solid, where the photons acquire an effective mass $(e^2n/m)^{1/2}$, n: density of the electrons.

Furthermore, in the Landau–Ginsburg theory for superconductivity [e.g. Fetter and Walecka, 1971, ch. 13] the spatial part $\boldsymbol{j}$ of the stationary "supercurrent" of the pair condensate coupled to the photon has the form

$$\boldsymbol{j} = -\frac{en}{m}\,\mathrm{grad}\,\theta - (ne^2/m)\,\boldsymbol{A}\,, \tag{8,51}$$

if one writes for the wave function $\psi(x) = (n(x))^{1/2}\,e^{i\theta(x)}$. The second term $\boldsymbol{j}_A = -(ne^2/m)\,\boldsymbol{A}$ on the r.h. side of eq. (8,51) is responsible for the Meissner effect ("Higgs mechanism") which gives the photon in the superconductor an effective mass $(ne^2/m)^{1/2}$. Comparing the current (8,49) with the expressions (8,50) and (8,51) – the interpretation of which is settled – there can hardly be any doubt concerning the interpretation of the term $j^\mu_A = -e^2\bar{u}u\,A^\mu/m$: it represents an induced effective mass of the photon!

This raises the following intriguing question concerning the Higgs mechanism in nonabelian gauge theories: In order to realize this symmetry-breaking mechanism, one has to introduce scalar ("Higgs"-) fields, the physical significance of which is highly controversial [e.g. Okun, 1981]. Perhaps it is possible to implement the symmetry-breaking, conventionally associated with the Higgs mechanism, by starting with a classical solution of a pure nonabelian gauge theory which has $\boldsymbol{E}^2 = \boldsymbol{B}^2$, and then putting the fermions in such a background field which in turn, acquires a (symmetry-breaking) effective mass by this interaction.

Bringing electric charges into a system which has $\boldsymbol{E}^2 = \boldsymbol{B}^2$ in general will drive the system into the electric "phase" $\boldsymbol{E}^2 > \boldsymbol{B}^2$. On the other hand, in chromodynamics the vacuum expection value

$$\langle 0|G_{\mu\nu}G^{\mu\nu}|0\rangle = \tfrac{1}{2}\langle 0|(\boldsymbol{B}^2 - \boldsymbol{E}^2)|0\rangle$$

of the gluon "condensate" can be estimated to be positive [Shifman et al., 1979], i.e., this system is in a magnetic "phase", in accordance with the heuristic picture that the QCD ground state is associated with the "condensation of magnetic monopoles" ['t Hooft, 1976; Mandelstam, 1976].

8.5. The property $L = 0$ in eulidean field theories and statistical mechanics

Many systems in statistical mechanics in states with long-range fluctuations, i.e. in the neighborhood of phase transitions – can be approximated by a nonlinear euclidean field theory [Landau, 1937; Patashinskij and Pokrovskij, 1964; Langer, 1965; Wilson, 1971; Moore, 1972; Wilson and Kogut, 1974, ch. 10; Wegner, 1976; Brézin et al., 1976; Amit, 1978; Glimm and Jaffe, 1981].

In d dimensions a popular Lagrangian for an n-component field $\boldsymbol{\varphi} = (\varphi^1, \ldots, \varphi^n)$ is given by

$$L = \tfrac{1}{2}\partial_j\boldsymbol{\varphi}\cdot\partial_j\boldsymbol{\varphi} + \tfrac{1}{2}\mu^2\,\boldsymbol{\varphi}\cdot\boldsymbol{\varphi} + \tfrac{1}{4}\lambda\,(\boldsymbol{\varphi}\cdot\boldsymbol{\varphi})^2\,, \tag{8,52}$$

where $\lambda > 0$, but sign and value of μ^2 is a function of the temperature T.

In statistical mechanics the Lagrangian L is the density of the free energy F [see, e.g., Wegner, 1976]. The field equations associated with the Lagrangian (8,52) are

$$\Delta\varphi^a = \mu^2\varphi^a + \lambda(\boldsymbol{\varphi}\cdot\boldsymbol{\varphi})\,\varphi^a\,, \qquad a = 1, \ldots, n\,. \tag{8,53}$$

It is evident from the expression (8,52) that for $\mu^2 \geq 0$ there are no solutions $\varphi \neq 0$ of the field equations with $L = 0$. Thus, we need spontaneous symmetry breaking in order to have such solutions: If $\mu^2 < 0$, then $\varphi_0 = (0, \ldots, 0, \varphi_0)$, $\varphi_0 = (-\mu^2/\lambda)^{1/2}$ can be chosen to be the ground state of the system. It has $L(\varphi_0) = -\frac{1}{4}\mu^4/\lambda$. Multiplying eqs. (8,53) by φ^a, summing over a and observing that $\partial_j\varphi\partial_j\varphi = \frac{1}{2}\Delta(\varphi)^2 - \varphi^a\Delta\varphi^a$, we see that for any solution of the field equations we have

$$L = \tfrac{1}{4}[\Delta\varphi^2 - \lambda(\varphi\cdot\varphi)^2]\,. \tag{8,54}$$

We define $\varphi_\perp = (\varphi^1, \ldots, \varphi^{n-1}, 0)$ and assume that there are approximate solutions $\varphi_\perp$ and $\varphi^n = \varphi_0 + \chi$ of the field eqs. (8,53) and the first order condition $L = 0$ such that

$$|\varphi_\perp| = [(\varphi^1)^2 + \cdots + (\varphi^{n-1})^2]^{1/2} \ll \varphi_0\,, \tag{8,55a}$$

$$|\chi| \ll |\varphi_\perp|\,. \tag{8,55b}$$

Condition (8,55a) implies, according to eq. (8,54),

$$\Delta\varphi_\perp^2 = \lambda\varphi_0^4 = \mu^4/\lambda\,. \tag{8,56}$$

Suppose that $\varphi_\perp$ depends only on $p \leq d$ of the variables x^i, $i = 1, \ldots, d$, say the p first ones. Then

$$\varphi_\perp^2 = A^2\rho_p^2\,, \qquad A^2 = \mu^4/(2p\lambda)\,, \qquad \rho_p = [(x^1)^2 + \cdots + (x^p)^2]^{1/2}\,, \tag{8,57}$$

is a solution of eq. (8,56). The condition (8,55a) here means

$$(-\mu^2/2p)^{1/2}\rho_p \ll 1\,, \tag{8,58}$$

that is, ρ_p has to be smaller than the correlation length $\xi \sim (-\mu^2)^{-1/2}$. Inserting the values for φ_0 and $\varphi_\perp^2$ into eqs. (8,53) we get

$$\Delta_p\varphi_\perp = \frac{\mu^4}{2p}\rho_p^2\varphi_\perp\,, \tag{8,59}$$

where Δ_p is the Laplace operator in p dimensions. Introducing hyperspherical coordinates ρ_p, ϕ, $\theta_1, \ldots, \theta_{p-2}$ in R^p [Erdélyi et al., vol. II, 1953, ch. XI] we have

$$\Delta_p = \frac{\partial^2}{\partial\rho_p^2} + \frac{p-1}{\rho_p}\frac{\partial}{\partial\rho_p} + \frac{1}{\rho_p^2}\mathrm{D}_p^{(2)}(\phi, \theta)\,, \tag{8,60}$$

where $\mathrm{D}_p^{(2)}$ is a second order differential operator in the variables ϕ, $\theta_1, \ldots, \theta_{p-2}$. The ansatz $\varphi_\perp = \rho_p\boldsymbol{A}_\perp(\phi, \theta)$ for a solution of eq. (8,59) leads to

$$(p-1)\boldsymbol{A}_\perp + \mathrm{D}_p^{(2)}\boldsymbol{A}_\perp = \frac{\mu^4}{2p}\rho_p^4\boldsymbol{A}_\perp\,. \tag{8,61}$$

In the region where our approximation (8,58) holds, we can neglect the r.h. side of eq. (8,61) and we see that the "transversal" components $\varphi_\perp^b$, $b = 1, \ldots, n-1$, are just the Goldstone modes: $\Delta_p \varphi_\perp^b = 0$. Therefore the functions $A_\perp^b$ can be chosen to be surface harmonics $Y_k(\phi, \theta)$ of degree k [see Erdélyi et al., l.c.]. Because

$$D_p^{(2)} Y_k = -k(k+p-2)\, Y_k\,,$$

we have the consistency condition $k^2 + k(p-2) = p-1$, with the 2 solutions

$$k = \begin{cases} 1 \\ 1-p \end{cases}. \tag{8,62}$$

The value $1-p$ is only possible for $p = 2$. The surface harmonics $Y_1(\phi, \theta)$ are essentially the p linearly independent and orthogonal unit vectors in R^p. As to the applications we consider two cases:

(i) Assume $p = n-1 = d-1$ and the field $\varphi_\perp$ to transform as a vector field with respect to the group $\mathrm{O}(n-1)$ – the field φ transforms according to $\mathrm{O}(n)$ –. In this case the fields form "vortices" the axis of which is parallel to the x^d-direction. It seems physically plausible to assume that the radius of the vortices are of the order of the correlation length ξ. Let $R(\alpha)$ be a fixed element of $\mathrm{O}(n-1)$. Then we can put

$$\varphi_\perp^b = A\,(R(\alpha)\,\boldsymbol{x})^b\,, \qquad \boldsymbol{x} \in \mathrm{R}^{n-1}\,, \tag{8,63}$$

which obeys eq. (8,57).

Example: ferromagnet in 3 dimensions [Brézin, 1976].

Here φ is the magnetization $\boldsymbol{M} = (M^1, M^2, M^3)$, $\varphi_0^3 = M_0$ is the ground state magnetization and $\mu^2 = a_0\tau$, $a_0 > 0$, $\tau = (T - T_c)/T_c$, where T_c is the critical temperature.

For the "small" transversal excitations we here have

$$\boldsymbol{M}_\perp = A(x^1 \cos\alpha - x^2 \sin\alpha,\, x^1 \sin\alpha + x^2 \cos\alpha,\, 0)\,.$$

From this we obtain for the "molecular" current density: $\boldsymbol{j} = \operatorname{curl} \boldsymbol{M} = 2A \sin\alpha\, \boldsymbol{e}_3$. Thus a constant current flows in 3-direction! Further remarks concerning the physics of this model can be found in* [Kastrup, 1981].

(ii) Assume that the components $\varphi_\perp^b$ transform according to some internal symmetry group $\mathrm{O}(n-1)$, not related to the group $\mathrm{O}(d)$ of the underlying euclidean space R^d. Then, because of eq. (8,57), the ansatz

$$\varphi_\perp^b = \sum_{i=1}^{p} C_i^b x^i$$

requires

$$\sum_b \sum_{i,j=1}^{p} C_i^b C_j^b x^i x^j = A^2 \rho_p^2\,,$$

* The discussion of the electric Meissner effect at the end of that paper is obviously incomplete.

which implies

$$\sum_b C_i^b C_j^b = A^2 \delta_{ij}$$

and we see that the matrix $(C_i^b)/A$ should be an orthogonal one, i.e. the index b has to take p different values, too. If $p < n - 1$, this can happen if some of the $\varphi_\perp^b$ vanish. If, on the other hand, $d > n - 1$, then one has to find a p such that $p = n - 1$.

We still have to determine the correction $\chi = \varphi^n - \varphi_0$ to the ground state $\varphi_0 = (0, \ldots, \varphi_0)$. Under the assumption (8,55b) we get from eq. (8,53)

$$\Delta\chi = \lambda\varphi_0(2\chi\varphi_0 + \varphi_\perp^2) .$$

Suppose further that $2\chi\varphi_0 \ll \varphi_\perp^2$. Then we get for χ the differential equation

$$\Delta\chi = \lambda\varphi_0\varphi_\perp^2 = \varphi_0(\mu^4/2p)\rho_p^2 , \tag{8,64}$$

which has the (special) solution

$$\chi^{(i)} = \frac{\mu^4}{8p(2+p)} \varphi_0 \rho_p^4 . \tag{8,65}$$

It follows immediately from the inequality (8,58) that the conditions $2\chi\varphi_0 \ll \varphi_\perp^2$ and $\chi \ll |\varphi_\perp|$ are indeed fulfilled!

There is a nice physical interpretation of the condition $L = 0$ in statistical mechanics: It was already mentioned above that in statistical mechanics the Lagrangian L equals the free energy density $f(x) = u(x) - Ts(x)$, where $u(x)$ is the density of the internal energy U and $s(x)$ the density of the entropy S. For dimensions $d \underset{(-)}{>} 2$ at very low temperatures in general the entropy term TS in $F = U - TS$ is considerably smaller than the internal energy U and the system can be in an "ordered" phase. With increasing temperature the entropy term increases and finally dominates, i.e. the system goes into an "unordered" phase. An approximate measure for this phase transition is the temperature T_c for which $U_c = T_c S_c$, i.e. for which $F_c = 0$. Thus, in analogy to the bifurcation phenomena in mechanics discussed above, the existence of "extremals" with $L = 0$ indicates the neighborhood of a "phase transition"!

The relation $F_c = 0$ has been used with considerable success in the statistical mechanics of phase transitions [e.g. Byckling, 1965; Felderhof, 1970; Kosterlitz and Thouless, 1973; Kosterlitz, 1974; Simon and Sokal, 1981 (with refs. to earlier papers)] and in euclidean lattice gauge theories [Banks et al., 1977; 't Hooft, 1978; Yonega, 1978; Kogut, 1979; Mack and Petkova, 1979, 1980]. The generalization $L = 0$ may provide a fruitful generalization for the analysis of related problems.

8.6. Bibliographical notes

Most of the relevant literature has already been mentioned in the preceding text. I got interested in solutions of the field equations with $L = 0$ by the (obvious) observation that many relations in Carathéodory's canonical theory become singular if $L = 0$: See, for instance, eqs. (5,4), (5,7), (5,10), (5,18), (5,40) and many others. The special significance of solutions of the field equations with $L = 0$

becomes evident only in Carathéodory's canonical theory, not, for instance, in the one of DeDonder–Weyl. The reason is, again, the fact that only in Carathéodory's canonical theory the transversal wave fronts are n-dimensional.

Singularities in transversality relations play an important role in catastrophe theory [see, e.g. Poston and Stewart, 1978 (with many refs.)]. Discussions of bifurcations in physics may be found in [Gurel and Rössler (eds.), 1979] and [Bardos and Bessis (eds.), 1979].

9. Unsolved problems and outlook

The material of the preceding chapters contains far more questions than answers as far as "hard" physical applications are concerned. Let me, therefore, list in a separate discussion some of those unsolved problems which appear to be relevant and which might be dealt with successfully in the near future:

1. In section 2.5 we saw that the propagation of classical wave fronts is generated by the function $F(t, q, k) = H^{-1}$, where the "wave vector" (k_j) is related to the canonical momenta p_j by $k_j = p_j/H$. It may be worthwhile to see how a quantum theory of such wave fronts looks like! This might be of special interest with respect to a quantization of field theories in the framework of Carathéodory's canonical theory (see nr. 7 below).

2. In mechanics the HJ wave fronts are the n-dimensional integral manifolds of the canonical 1-form $\theta = -H\,\mathrm{d}t + p_j\,\mathrm{d}q^j$ on R^{1+n}, whereas the 1-dimensional extremals are the "characteristic" integral manifolds of the 2-form $\mathrm{d}\theta$ on R^{1+2n} with coordinates (t, q, p). In addition, the wave fronts can be generated by an n-parameter family ("field") of characteristics. These properties may also be formulated in terms of "relative" and "absolute integral invariants" [E. Cartan, 1922; Godbillon, 1969, ch. VIII; Choquet-Bruhat et al., 1977, ch. IV, C].

The situation is quite different for field theories, as we have seen in section 3.3: The wave fronts are again integral manifolds associated with the fundamental canonical form Ω. However, their dimension now depends crucially on the rank of Ω. Furthermore, the 2-dimensional extremals Σ_0^2 are integral manifolds of the "variational" system $\mathrm{I}[i(\partial_\mu)\,\mathrm{d}\Omega,\ i(\partial_a)\,\mathrm{d}\Omega,\ i(\partial/\partial p_a^\mu)\,\mathrm{d}\Omega]$, or the "proper" variational system $\mathrm{I}[\omega^a,\ \mathrm{d}\omega^a,\ i(\partial_a)\,\mathrm{d}\Omega]$, whereas, at least in the case of two independent variables, the dimension of the characteristic integral manifolds of the form $\mathrm{d}\Omega$ depends on the rank of the matrix $(\partial p_a^\mu/\partial v_\nu^b)$: For $\Omega = \Omega_0$ (DeDonder–Weyl) the rank c of $\mathrm{d}\Omega_0$ is $c = n + 2 + \mathrm{rank}\ (\partial\pi_a^\mu/\partial v_\nu^b)$ [von Rieth and Kastrup, 1983]. Thus, if $(\partial\pi_a^\mu/\partial v_\nu^b)$ is regular, the characteristic manifolds of $\mathrm{d}\Omega_0$ are 0-dimensional!

On the other hand, for gauge theories, where $(\partial\pi_a^\mu/\partial v_\nu^b)$ is singular, the characteristic manifolds of $\mathrm{d}\Omega_0$ have dimension >0 (e.g. for E-dynamics in 2 dimensions they are given by x^μ = const., z^μ = const., $\mu = 0,1$, $v_0^1 + v_1^0$ = const., i.e. the characteristic manifolds are 3-dimensional, with "running" variables v_0^0, v_1^1 and $v_0^1 - v_1^0$).

We have seen in chapter 3, too, that the problems concerning the involutiveness of the variational systems are nontrivial. Furthermore, the question arises, how to generalize the concept of integral invariants to field theories [DeDonder, 1935; Dedecker, 1977a] and what use can be made of them.

All these problems need additional investigation.

3. One of the main problems we encountered in the context of HJ theories of fields is: How to solve the integrability conditions (4,6) or (5,71). This problem may be rephrased in the following way, starting again from mechanics:

Suppose, through each point $(t, q) \in \mathrm{G}^{1+n}$ passes just one extremal $q(t)$. Such a family of extremals can be parametrized by n variables u^j, $j = 1, \dots, n$. Then, in terms of modern mathematics [Lawson Jr., 1977, with many refs.], the region G^{1+n} is "foliated" by 1-dimensional "leaves" $q^j(t) = f^j(t; u)$. If the Lagrangian brackets $[u^j, u^k]$ vanish, then the 1-dimensional leaves generate transversal n-dimensional leaves, the wave fronts (see section 2.2). The situation is more complicated in field theories: Suppose, we have found an n-parameter family of extremals $z^a = f^a(x; u)$, such that through each point (x, z) G^{m+n} passes just one extremal, then we have a foliation of G^{m+n} by m-dimensional leaves (the extremals form m-dimensional submanifolds in G^{m+n}!). However, we have seen in section 4.4 that such a foliation in general will not generate n-dimensional transversal leaves (wave fronts), because the associated canonical form $\tilde{\Omega}$ in general will have a rank larger than m.

On the other hand, given n-dimensional wave fronts $S^\mu(x, z) = \sigma^\mu = \text{const.}$, $\mu = 1, \dots, m$, in G^{m+n}, where the functions S^μ are solutions of a HJ equation associated with a canonical m-form Ω, then those wave fronts are the n-dimensional leaves of a foliation of G^{m+n}. However, these leaves in general will not generate m-dimensional transversal leaves which are the extremals of the variational system of Ω, because the integrability conditions (4,6) or (5,71) in general will not be satisfied. But they may generate transversal submanifolds (leaves) of dimension less than m, which, in turn, may be helpful for the construction of the wave fronts.

It is obvious that the integrability problems encountered in the chapters 4 and 5 are part of this subject.

4. Of special interest are the HJ currents, associated with each parameter an extremal depends upon. It may turn out that in many cases the conserved quantities generated by such currents will not be very useful. However, only a detailed investigation of specific models can tell whether the classical HJ currents generate a structure which may be helpful for the quantization of the system.

5. Another interesting question is, whether the concept of a "complete" integral as discussed in sections 4.3 and 5.6 can be developed into a workable tool for solving the Euler–Lagrange field equations.

6. I mentioned already in the Introduction that field theories, when analyzed from a "mechanical" point of view, are systems with an infinite number of degrees of freedom, the time evolution of which takes place in an infinite dimensional phase space [Chernoff and Marsden, 1974; Abraham and Marsden, 1978, ch. 5; Itzykson and Zuber, 1980, chs. 1 and 3]. That approach, which – justly so – is the prevailing one, is to some extent complementary to the more geometrical interpretation adopted in the preceding chapters. Thus, the question arises, how these two points of view are interrelated and how they can profit from each other! That relation is important for the corresponding quantum theory, too.

7. *That* problem which is probably the most important one concerning far-reaching physical applications of the more general canonical framework discussed above, has hardly been mentioned at all in the text. Here certainly one will have to deal with difficult problems:

For a quantum field theory not only those configurations of the field variables are important for which the "classical" relations $\omega^a = dz^a - v^a_\mu \, dx^\mu = 0$ hold, but the "off-shell" configurations are essential, too! However, we modify the off-shell properties of the fields, if we define the canonical momenta p_a according to eq. (3,9) with $h_{ab} \neq 0$. The problems one has to deal with may be illustrated by two examples:

(i) Suppose we have n (coupled) real scalar fields $z^a = \varphi^a(x)$, $a = 1, \dots, n$, $x = (x^0, x^1)$, in a 2-dimensional Minkowski space. Then we have $\pi^0_a = \partial_0 \varphi^a$, $\pi^1_a = -\partial_1 \varphi^a$ and the conventional "DeDonder–Weyl" symplectic structure of the field configurations is determined by the Poisson brackets [e.g. Itzykson and Zuber, 1980, ch. 1]:

$$\{\varphi_a(x), \varphi_b(y)\}_{x^0=y^0} = 0\,, \tag{9,1a}$$

$$\{\pi^0_a(x), \varphi^b(y)\}_{x^0=y^0} = \delta^b_a \delta(x^1 - y^1)\,, \tag{9,1b}$$

$$\{\pi^0_a(x), \pi^0_b(y)\}_{x^0=y^0} = 0\,. \tag{9,1c}$$

The basic relations (9,1) imply

$$\{\varphi^a(x), \pi^1_b(y)\}_{x^0=y^0} = 0\,, \tag{9,2a}$$

$$\{\pi^0_a(x), \pi^1_b(y)\}_{x^0=y^0} = \delta_{ab}\partial^{(x)}_1 \delta(x^1 - y^1)\,, \tag{9,2b}$$

$$\{\pi^1_a(x), \pi^1_b(y)\}_{x^0=y^0} = 0\,. \tag{9,2c}$$

Assume now that the quantities h_{ab} are "external" fields, that is to say, they are functions of x only – e.g. constants – and do not depend on the quantities φ^a and π^μ_a. Then the relations (9,1) imply for $p^\mu_a = \pi^\mu_a - \varepsilon^{\mu\nu} h_{ab} v^b_\nu$:

$$\{p^0_a(x), \varphi^b(y)\}_{x^0=y^0} = \delta^b_a\, \delta(x^1 - y^1)\,, \tag{9,3a}$$

$$\{p^0_a(x), p^0_b(y)\}_{x^0=y^0} = 0\,, \tag{9,3b}$$

$$\{p^0_a(x), p^1_b(y)\}_{x^0=y^0} = \Big(\delta_{ab} + \sum_c h_{ac} h_{bc}\Big)\, \partial^{(x)}_1\, \delta(x^1 - y^1)\,, \tag{9,3c}$$

$$\{p^1_a(x), \varphi^b(y)\}_{x^0=y^0} = h_{ab}\delta(x^1 - y^1)\,, \tag{9,3d}$$

$$\{p^1_a(x), p^1_b(y)\}_{x^0=y^0} = 0\,. \tag{9,3e}$$

We see: If we replace the momenta π^μ_a by the quantities p^μ_a, then the basic relations (9,1) remain unchanged, but the relations (9,2a,b) are replaced by new ones. Thus, the symplectic structure – if there is one at all – of the variables φ^a, p^μ_b is different from the conventional one with respect to the variables φ^a and π^μ_b!

(ii) Carathéodory's canonical theory as discussed in chapter 5. The rather complicated relation between the "velocities" v^a_μ and the canonical momenta p^μ_a at first sight seems to raise trouble if one compares it with the conventional canonical quantization procedure which used the momenta π^μ_a. On the other hand, it is certainly very intriguing that this conventional (DeDonder–Weyl) canonical formalism appears to be the zero order approximation if one expands Carathéodory's canonical quantities p^μ_a and H_c in powers of L^{-1}! Furthermore, we have seen that the canonical E. Hölder transformation (section 5.7) casts a given field theory into a "mechanical" canonical form on the hypersurfaces $S^{\bar\mu}(x, z) = \sigma^{\bar\mu} = \text{const.}$, $\bar\mu = 2, \ldots, m$. This suggests to introduce for the quantities $\hat z^a = \hat z^a(y^1, \boldsymbol{\sigma}_2)$, $p_a = \hat p^1_a(x^1, \boldsymbol{\sigma}_1)$, with x^1, y^1 as time-coordinates and $\boldsymbol{\sigma} = (\sigma^2, \ldots, \sigma^m)$, the Poisson brackets

$$\{\hat p^1_a(x^1, \boldsymbol{\sigma}_1), \hat z^b(y^1, \boldsymbol{\sigma}_2)\}_{x^1=y^1} = \delta^b_a\, \delta(\boldsymbol{\sigma}_1 - \boldsymbol{\sigma}_2) \tag{9,4}$$

with respect to the surfaces $S^{\bar\mu}(x, z) = \sigma^{\bar\mu} = \text{const.}$ However, since these surfaces are dynamical ones which depend on the given Hamilton function $H_c(x, z, p)$, such a postulate might not be compatible with the conventional "kinematical" Poisson brackets (9,1). In addition: Quantization means that the variables z^a become operator-valued distributions, whereas the coordinates x remain c-numbers. As the E. Hölder transformation "mixes" both quantities, one certainly has conceptual problems for a "quantized" version of this transformation.

Another perhaps more promising approach to the problem of quantization may be a *phase space* path integral formalism in the E. Hölder frame. Canonical transformations in phase-space path integrals are, however, full of pitfalls, too [Gervais and Jevicki, 1976].

Despite all these difficulties: I have the strong impression that Carathéodory's canonical theory bears the possibility for a qualitatively new approach to the quantization of field theories!

Another problem, to be dealt with in the context of quantization is the following: In quantum mechanics solutions of the HJ equation are useful for WKB-approximations. Similarly, solutions of the HJ equations for field theories may be useful for semiclassical approximations of path-integrals for those fields.

8. We have seen in section 3.4 that the use of forms $\omega^a = dz^a - v^a_\mu\, dx^\mu$ and of the ideal $I[\omega^a]$ generated by them, provides a natural framework for the description of gauge invariance in E-dynamics. A generalization to nonabelian gauge theories appears desirable and will probably have to use covariant differentials [see, e.g., Pham Mau Quan, 1969, ch. V.21; Eguchi et al., 1980, ch. 5]:

Given a connection Γ with coefficients $\Gamma^a_{\mu b}$ we can decompose the tangent vectors $X_\mu = \partial_\mu + v^a_\mu \partial_a$ into horizontal and vertical components:

$$X_\mu = D_\mu + \bar{v}^a_\mu \partial_a\,, \qquad D_\mu = \partial_\mu - \Gamma^a_{\mu b}\, z^b \partial_a\,, \qquad \bar{v}^a_\mu = v^a_\mu + \Gamma^a_{\mu b} z^b\,.$$

If $\sigma^a = dz^a + \Gamma^a_{\mu b} z^b\, dx^\mu$ is the connection 1-form of Γ, then $\omega^a = \sigma^a - \bar{v}^a_\mu\, dx^\mu$ etc..

9. The rather uncomplicated "translation" of the Lagrangian $L = \frac{1}{2}[(v_1)^2 (v_2)^2 - (v_1 \cdot v_2)^2]$ for the relativistic string – eq. (5,55) – into Carathéodory's formalism [Kastrup and Rinke, 1981], suggests a similar approach to nonlinear σ-models e.g., with Lagrangians $L = \frac{1}{4}[(\partial_\mu \varphi^a)^2 (\partial_\nu \varphi^b)^2 - (\partial_\mu \varphi^a \partial_\nu \varphi^a)^2]$, where the fields have to obey the constraint $\varphi^a \varphi^a = 1$. (As to the literature on nonlinear σ-models see the articles by Maison [1979], Fröhlich [1980] and Kafiev [1981].)

In addition, the expression $A_\mu = \frac{1}{2}i(z_a \partial_\mu z_a - (\partial_\mu z_a) z_a)$ for composite gauge fields in CP^{n-1}-models [D'Adda et al., 1978] bears some resemblance to the formula $A_\mu = \lambda S^2 \partial_\mu S^1$, which relates the "wave fronts" associated with a relativistic string to the potential A_μ of an electromagnetic field of rank 2 [Kastrup and Rinke, 1981].

10. The "bifurcation" condition $L = 0$ (or $L = \text{const.}$) for mechanical systems has to be analyzed further, especially for many-particle systems. The relation between the classical bifurcation curves $E = E_0(l)$ and the corresponding quantum mechanical energy levels has to be clarified.

Of special interest is to find "nonabelian" solutions of Yang–Mills theories in Minkowski space with the property $\boldsymbol{E}^2 = \boldsymbol{B}^2$, calculate the quantum fluctuations "around" them and put spinor fields into such "background" fields.

The condition $H_c = 0$, or $|(T^\mu_\nu)| = 0$, for fields in general has to be analyzed, too.

11. Throughout the paper we have only dealt with differential forms on "normal" manifolds. However, in recent years "supermanifolds", with differential forms "graded" into "bosonic" and "fermionic" ones, have become quite popular in theoretical physics [Zumino, 1976; Sternberg, 1977; Kostant, 1977; Hermann, 1977; Salam and Strathdee, 1978; Berezin 1979a, b]. Here the question arises, how the concepts discussed in this review can be generalized to such supermanifolds and what use can be made of such a framework?

Obviously there are more open problems! The above list indicates that – at the very least – the more general canonical framework discussed in the previous chapters allows one to formulate a number of interesting questions, the answers of which may provide important new insights into the rich structure and content of dynamical systems in physics.

Acknowledgements

I am indebted to M. Rinke and J. von Rieth for a critical reading of several chapters and I am grateful to K.-P. Dörpelkus and A. Nüsser for critical questions concerning an earlier version of the manuscript. Parts of the final improvements on this review were written during September 1982 at the Max-Planck-Institut in Munich and in February 1983 at CERN. I thank both institutions, especially W. Zimmermann and M. Jacob, for their kind hospitality.

References

Abraham, R. and J.E. Marsden (1978), Foundations of Mechanics, 2nd ed. (Benjamin/Cummings Publ. Co., Reading, Mass.).
Adler, S.L. (1970), Perturbation theory anomalies, in: Lectures on Elementary Particles and Quantum Field Theory, vol. I, eds. St. Deser, M. Grisaru and H. Pendleton (M.I.T. Press, Cambridge, Mass.) pp. 1–164.
Aharonov, Y. and D. Bohm (1959), Significance of electromagnetic potentials in the quantum theory, Phys. Rev. 115, 485–491.
Aldaya, V. and J.A. de Azcárraga (1980), Geometric formulation of classical mechanics and field theory, Riv. Nuov. Cim. 3, N. 10.
Amit, D.J. (1978), Field Theory, the Renormalization Group, and Critical Phenomena (McGraw-Hill Book Co., N.Y.).
Anderson, I.M. and T. Duchamp (1980), On the existence of global variational principles, Amer. J. Math. 102, 781–868.
Anderson, P.W. (1963), Plasmons, gauge invariance, and mass, Phys. Rev. 130, 439–442.
Arnold, V.I. (1967), On a characteristic class entering in quantization conditions, Funct. Anal. Appl. 1, 1–13.
Arnold, V.I. (1978), Mathematical Methods of Classical Mechanics (Springer-Verlag, N.Y., Heidelberg, Berlin).
Banks, T., R. Myerson and J. Kogut (1977), Phase transitions in abelian gauge theories, Nucl. Phys. B129, 493–510.
Bardos, C. and D. Bessis (1979), eds., Bifurcation Phenomena in Mathematical Physics and Related Topics (D. Reidel Publ. Co., Dordrecht).
Barthel, W. (1958), Über homogene Funktionen auf dem Graßmann-Kegel, Archiv Math. 9, 262–274.
Bashilov, Yu.A. and S.V. Pokrovsky (1978), Quantum fluctuations in the vicinity of an instanton in the SU(N) group, Nucl. Phys. B143, 431–444.
Beckert, H. (1974), Einige Bemerkungen zur lokalen Existenztheorie in der mehrdimensionalen Variationsrechnung, Beitr. z. Anal. 6, 17–23.
Belavin, A.A., A.M. Polyakov, A.S. Schwartz and Yu.S. Tyupkin (1975), Pseudo particle solutions of the Yang–Mills equations, Phys. Lett. 59 B, 85–87.
Bellman, R. (1970), Introduction to Matrix Analysis, 2nd ed. (McGraw-Hill Book Co., N.Y.).
Berestetskij, V.B., E.M. Lifshitz and L.P. Pitaevskij (1971), Relativistic Quantum Theory, vol. 4, I of the "Course of Theoretical Physics" by Landau and Lifshitz (Pergamon Press, Oxford).
Berezin, F.A. (1979a), The mathematical basis of supersymmetric field theories, Sov. J. Nucl. Phys. 29, 857–866.
Berezin, F.A. (1979b), Differential forms on supermanifolds, Sov. J. Nucl. Phys. 30, 605–609.
Berry, M.V. and K.E. Mount (1972), Semiclassical approximations in wave mechanics, Reports Prog. Phys. 35, 315–397.
Berry, M. (1981), Singularities in waves and rays, in: Physics of Defects, eds. R. Balian, M. Kléman and J.-P. Poirier (North-Holland Publ. Co., Amsterdam) pp. 453–543.
Biran, B. (1980), Carathéodory's canonical field equations, Acad. Roy. Belg. Bull. Cl. Science (5e Sér.) 66, 143–148.
Bjorken, J.D. and S.D. Drell (1964), Relativistic Quantum Mechanics (McGraw-Hill Book Co., N.Y.).
Bjorken, J.D. and S.D. Drell (1965), Relativistic Quantum Fields (McGraw-Hill Book Co., N.Y.).
Boerner, H. (1936), Über die Extremalen und geodätischen Felder in der Variationsrechnung der mehrfachen Integrale, Math. Ann. 112, 187–220.
Boerner, H. (1940a), Variationsrechnung aus dem Stokesschen Satz, Mathem. Zeitschr. 46, 709–719.
Boerner, H. (1940b), Über die Legendresche Bedingung und die Feldtheorien in der Variationsrechnung der mehrfachen Integrale, Mathem. Zeitschr. 46, 720–742.
Boerner, H. (1953), Carathéodory's Eingang zur Variationsrechnung, Jahresber. Deutsch. Math.-Verein. 56, 31–58.
Bogoliubov, N.N. and D.V. Shirkov (1959), Introduction to the Theory of Quantized Fields (Interscience Publ., N.Y.).
Born, M. (1934), On the quantum theory of the electromagnetic field, Proc. Roy. Soc. A143, 410–434.
Born, M. and E. Wolf (1975), Principles of Optics, 5th ed. (Pergamon Press, Oxford).

Brézin, E. (1976), Applications of the renormalization group to critical phenonena, in: Methods in Field Theory, eds. R. Balian and J. Zinn-Justin (North-Holland Publ. Co., Amsterdam) pp. 329–386.

Brézin, E., J.C. Le Guillou and J. Zinn-Justin (1976), Field theoretical approach to critical phenomena, in: Phase Transitions and Critical Phenomena, vol. 6, eds. C. Domb and M.S. Green (Academic Press, London, N.Y., San Francisco) pp. 125–247.

Brown, L.S., R.D. Carlitz and C. Lee (1977), Massless excitations in pseudoparticle fields, Phys. Rev. D16, 417–422.

Byckling, E. (1965), Vortex lines and the λ-transition, Ann. Phys. 32, 367–376.

Callan Jr., C.G., R.F. Dashen and D.J. Gross (1979), Semiclassical methods in quantum chromodynamics: toward a theory of hadron structure, in: Quantum Chromodynamics (AIP Conf. Proc. 55) eds. W. Frazen and F. Henyey (Am. Inst. Physics, N.Y.).

Carathéodory, C. (1929), Über die Variationsrechnung bei mehrfachen Integralen, Acta Sci. Math. (Szeged) 4, 193–216; reprinted in: C. Carathéodory, Gesamm. Mathem. Schriften, Bd. 1 (Beck'sche Verlagsbuchhandlung, München 1954) pp. 401–426.

Carathéodory, C. (1930), Variationsrechnung, in: Die Differential- und Integralgleichungen der Mechanik und Physik, Bd. 1, 2. Aufl., hrsg, von Ph. Frank und R. v. Mises (Vieweg Verl., Braunschweig) pp. 227–279.

Carathéodory, C. (1935), Variationsrechnung und Partielle Differentialgleichungen erster Ordnung (B.G. Teubner, Leipzig und Berlin), Engl. transl.: Calculus of Variations and Partial Differential Equations (Holden-Day, San Francisco, 1965).

Carathéodory, C. (1937), Geometrische Optik, Ergebn. Math. Grenzgeb. 4, 573–680.

Cartan, E. (1922), Leçons sur les Invariants Intégraux (Hermann, Paris).

Cartan, E. (1945), Les Systèmes Différentiels Extérieur et leurs Application Géométriques (Hermann, Paris).

Chernoff, P.R. and J.E. Marsden (1974), Properties of Infinite Dimensional Hamiltonian Systems (Lect. Notes in Math. 425; Springer-Verlag, Berlin, Heidelberg, N.Y.).

Choquet-Bruhat, Y., C. DeWitt-Morette and M. Dillard-Bleick (1977), Analysis, Manifolds and Physics (North-Holland Publ. Co., Amsterdam, N.Y., Oxford).

Clebsch, A. (1859), Ueber die zweite Variation vielfacher Integrale, Journ. f. reine u. angew. Math. 56, 122–148.

Coleman, S. (1977), Non-abelian plane waves, Phys. Lett. 70B, 59–60.

Coleman, S. (1979), The uses of instantons, in: The Whys of Subnuclear Physics, ed. A. Zichichi (Plenum Press, N.Y. and London) pp. 805–941.

Collins, P.A. and R.W. Tucker (1976), Classical and quantum mechanics of free relativistic membranes, Nucl. Phys. B112, 150–176.

D'Adda, A., M. Lüscher and P. Di Vecchia (1978), A $1/n$ expandable series of non-linear σ models with instantons, Nucl. Phys. B146, 63–76.

Dashen, R.F., B. Hasslacher and A. Neveu (1974, I), Nonperturbative methods and extended-hadron models in field theory. I. Semiclassical functional methods, Phys. Rev. D10, 4114–4129.

Dashen, R.F., B. Hasslacher and A. Neveu (1974, II), Nonperturbative methods and extended-hadron models in field theory. II. Two-dimensional models and extended hadrons, Phys. Rev. D10, 4130–4142.

Debever, R. (1937), Les champs de Mayer dans le calcul des variations des intégrales multiples, Acad. Roy. Belg. Bull. Cl. Sci. (5e Sér.) 23, 809–815.

Debever, R. (1941), Quelques conséquences de la condition nécessaire d'Hadamard du calcul des variations des intégrales multiple, Bull. Soc. Roy. Sci. Liège 10, 584–589.

Dedecker, P. (1950), Sur les intégrales multiple du calcul des variations, Comptes rendus du III^e Congrès National des Sciences, vol. 2 (Bruxelles) pp. 29–35.

Dedecker, P. (1951), Sur le théorème de la circulation de V. Bjerknes et la théorie des invariants intégraux, Miscell. de Institut Roy. Météorl. Belg., fasc. 36.

Dedecker, P. (1952a), Les systèmes d'équations extérieures, Colloque de Topologie de Strasbourg (notes polycopiées).

Dedecker, P. (1952b), Equations différentielles extérieures et calcul des variations, Colloque de Topologie des Strasbourg (notes polycopiées).

Dedecker, P. (1953), Calcul des variations, formes différentielles et champ géodésiques, en: Géométrie différentielle, Strasbourg 1953 (Colloques intern. du C.N.R.S., nr. 52, Paris) pp. 17–34.

Dedecker, P. (1957a), Calcul des variations et topologie algébrique, Mém. Soc. Roy. Sci. Liège, t. 19 (4e sér.), fasc. I.

Dedecker, P. (1957b), A property of differential forms in the calculus of variations, Pacific J. Math. 7, 1545–1549.

Dedecker, P. (1977a), On the generalization of symplectic geometry to multiple integrals in the calculus of variations, in: Differential Geometrical Methods in Mathematical Physics (Lecture Notes in Mathem. 570), eds. K. Bleuler and A. Reetz (Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 395–456.

Dedecker, P. (1977b), Généralisation d'une formule de H.A. Schwarz relative aux surfaces minima, C. R. Acad. Sc. Paris 285 (A), 23–26.

Dedecker, P. (1977c), Généralisation d'une formule de H.A. Schwarz aux intégrales multiples du calcul des variations, C. R. Acad. Sc. Paris 285 (A), 59–61.

Dedecker, P. (1977d), Intégrales complètes de l'équation aux dérivées partielles de Hamilton–Jacobi d'une intégrale multiple, C. R. Acad. Sc. Paris 285 (A), 123–126.

Dedecker, P. (1978), Problèmes variationnels dégénérés, C. R. Acad. Sc. Paris 286 (A), 547–550.

Dedecker, P. and W.M. Tulczyjew (1980), Spectral sequences and the inverse problem of the calculus of variations, in: Differential Geometrical Methods in Mathematical Physics, eds. P.L. García, A. Pérez-Rendón and J.M. Souriau (Lect. Notes in Math. 836, Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 498–503.

De Donder, Th. (1911), Sur les équations canoniques de Hamilton–Volterra, Acad. Roy. Belg. Cl. Sci. Mem. in 4°, t. 3, fasc. 4.

De Donder, Th. (1913), Sur le théorème d'independence de Hilbert, C. R. Acad. Sci. Paris 156, 868–870.

De Donder, Th. (1935), Theorie Invariantive du Calcul des Variations, Nouv. éd. (Gauthier-Villars, Paris).

De Witt-Morette, C., A. Maheshwari and B. Nelson (1979), Path integration in non-relativistic quantum mechanics, Phys. Reports 50, 255–372.

Dieudonné, J. (1972, 1974), Treatise on Analysis, vols. III, IV (Academic Press, N.Y. and London).
Dirac, P.A.M. (1933), Homogeneous variables in classical dynamics, Proc. Camb. Phil. Soc. 29, 389–400.
Dirac, P.A.M. (1964), Lectures on Quantum Mechanics (Yeshiva Univ., N.Y.).
Dixon, A.C. (1909), On the relation between Pfaff's problem and the calculus of variations, Proc. London Mathem. Soc. 7, 422–429.
Dokshitzer, Yu.L., D.I. Dyakonov and S.I. Troyan (1980), Hard processes in quantum chromodynamics, Phys. Reports 58, 269–395.
Duistermaat, J.J. (1974), Oscillatory integrals, Lagrange immersions and unfolding of singularities, Commun. Pure and Appl. Math. 27, 207–281.
Eguchi, T., P.B. Gilkey and A.J. Hanson (1980), Gravitation, gauge theories and differential geometry, Phys. Reports 66, 213–393.
Englert, F. and P. Windey (1978), Electric confinement and magnetic superconductors, Nucl. Phys. B135, 529–545.
Erdélyi, A., W. Magnus, F. Oberhettinger and F.G. Tricomi (1953), Higher Transcendental Functions, vol. II (McGraw-Hill Book Co., N.Y.).
Estabrook, F.B. and H.D. Wahlquist (1975), The geometric approach to sets of ordinary differential equations and Hamiltonian dynamics, SIAM Rev. 17, 201–220.
Estabrook, F.B. (1980), Differential geometry as a tool for applied mathematics, in: Geometric Approaches to Differential Equations, ed. R. Martin (Lect. Notes in Math. 810; Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 1–22.
Felderhof, B.U. (1970), Do phase transitions exist?, Nature 225, 17–20.
Fetter, A.L. and J.D. Walecka (1971), Quantum Theory of Many-Particle Systems (McGraw-Hill Book Co., N.Y.).
Foldy, L.L. and S.A. Wouthuysen (1950), On the Dirac theory of spin 1/2 particles and its non-relativistic limit, Phys. Rev. 78, 29–36.
Fréchet, M. (1905), Sur une extension de la méthode de Jacobi–Hamilton, Ann. Mat. Pura Appl. 11, 187–199.
Fröhlich, J. (1980), A new look at generalized, non-linear σ-models and Yang–Mills theory, in: Quantum Fields-Algebras, Processes, ed. L. Streit (Springer-Verlag, Wien, N.Y.) pp. 361–378.
Fuchs, G. (1973), Zur Feldtheorie Lagrangescher Variationsprobleme in Parameterdarstellung, Beitr. z. Anal. 5, 35–53.
Funk, P. (1970), Variationsrechnung und ihre Anwendung in Physik und Technik, 2. Aufl. (Springer-Verlag, Berlin, Heidelberg, N.Y.).
Gallissot, F. (1958), Les formes extérieures et la méchanique des milieux continus, Ann. Inst. Fourier (Grenoble) 8, 291–335.
García, P.L. and A. Pérez-Rendón (1969), Symplectic approach to the theory of quantized fields. I. Comm. Math. Phys. 13, 24–44.
García, P.L. and A. Pérez-Rendón (1971), Symplectic approach to the theory of quantized fields. II. Arch. Rational Mech. Anal. 43, 101–124.
García, P. (1974), The Poincaré–Cartan invariant in the calculus of variations, Symposia Mathematica, vol. 14 (Academic Press, N.Y.) pp. 219–246.
García, P.L. (1977), Reducibility of the symplectic structure of classical fields with gauge theory, in: Differential Geometrical Methods in Mathematical Physics, eds. K. Bleuler and A. Reetz (Lect. Notes in Math. 570; Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 365–376.
García, P.L. and A. Pérez-Rendón (1978), Reducibility of the symplectic structure of minimal interactions, in: Differential Geometrical Methods in Mathematical Physics II, eds. K. Bleuler, H.R. Petry and A. Reetz (Lect. Notes in Math. 676, Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 409–433.
García, P.L., A Pérez-Rendón and J.M. Souriau, eds. (1980), Differential Geometrical Methods in Mathematical Physics (Lect. Notes Math. 836; Springer-Verlag, Berlin, Heidelberg, N.Y.).
Gasiorowicz, St. (1966), Elementary Particle Physics (J. Wiley and Sons, N.Y., London, Sydney).
Gawędzki, K. (1972), On the geometrization of the canonical formalism in the classical field theory, Reports Math. Phys. 3, 307–326.
Gawędzki, K. and W. Kondracki (1974), Canonical formalism for the local-type functionals in the classical field theory, Reports Math. Phys. 6, 465–476.
Gelfand, I.M. and S.V. Fomin (1963), Calculus of Variations (Prentice-Hall, Englewood Cliffs).
Gelfand, I.M. and L.A. Dikii (1975), Asymptotic behaviour of the resolvent of Sturm–Liouville equations and the algebra of the Korteweg–de Vries equation, Russ. Math. Surveys 30, 77–113.
Gelfand, I.M. and L.A. Dikii (1977), The resolvent and Hamiltonian systems, Funct. Analysis and its Appl. 11, 93–105.
Gervais, J.-L. and A. Jevicki (1976), Point canonical transformations in the path integral, Nucl. Phys. B110, 93–112.
Glassey, R.T. and W.A. Strauss (1980), Propagation of the energy of Yang–Mills fields, in: Bifurcation Phenomena in Mathematical Physics and Related Topics, eds. C. Bardos and D. Bessis (D. Reidel Publ. Co., Dordrecht) pp. 231–241.
Glimm, J. and A. Jaffe (1981), Quantum Physics (Springer-Verlag, N.Y., Heidelberg, Berlin).
Godbillon, C. (1969), Géometrie différentielle et méchanique analytique (Hermann, Paris).
Goldschmidt, H. and S. Sternberg (1973), The Hamilton–Cartan formalism in the calculus of variations, Ann. Inst. Fourier (Grenoble) 23, 203–267.
Groß, W. (1916), Das isoperimetrische Problem bei Doppelintegralen, Monatsh. Math. Phys. 27, 70–120.
Guckenheimer, J. (1974), Caustics, in: Global Analysis and its Applications, vol. II (Intern. Atomic Energy Agency, Vienna) pp. 281–289.
Guillemin, V.W. and S. Sternberg (1977), Geometric Asymptotics (Math. surveys 14; Am. Math. Soc. Providence).
Gurel, O. and O.E. Rössler, eds. (1979), Bifurcation Theory and Application in Scientific Disciplines (Ann. N.Y. Acad. Sc. No. 316; N.Y. Acad. Sc.).
Hadamard, M. (1902), Sur une question de calcul des variations, Bull. Soc. Math. France 30, 253–256.
Hadamard, M. (1905), Sur quelques questions de calcul des variations, Bull. Soc. Math. France 33, 73–80.
Hermann, R. (1965), E. Cartan's geometric theory of partial differential equations, Advances in Math. 1, 265–317.
Hermann, R. (1977a), Differential Geometry and the Calculus of Variations, 2nd ed. (Interdisciplinary Mathematics XVII; Math. Sci. Press, Brookline).
Hermann, R. (1977b), Quantum and Fermion Differential Geometry A (Interdisciplinary Mathematics XVI; Math. Sci. Press, Brookline).
Hestenes, M.R. (1966), Calculus of Variations and Optimal Control Theory (J. Wiley and Sons, Inc., N.Y., London, Sydney).
Hilbert, D. (1900), Mathematische Probleme, Nachr. d. Königl. Ges. d. Wiss. Göttingen, Math.-Physik. Kl. 1900, 253–297; reprinted in: D. Hilbert, Gesamm. Abhandl., Bd. III (Springer-Verlag, Berlin, Heidelberg 1935) pp. 290–329.

Hilbert, D. (1906), Zur Variationsrechnung, Math. Ann. 62, 351–368.

Hill, E.L. (1951), Hamilton's principle and the conservation theorems of mathematical physics, Rev. Mod. Phys. 23, 253–260.

Hodge, W.V.D. and D. Pedoe (1953, 1952), Methods of Algebraic Geometry, vols. I, II (Cambridge Univ. Press).

Hölder, E. (1939), Die infinitesimalen Berührungstransformationen in der Variationsrechnung, Jahresber. Deutsch. Math.-Verein. 49, 162–178.

Hölder, E., R. Klötzler, S. Gähler und St. Hildebrandt (1966), Entwicklungslinien der Variationsrechnung seit Weierstraß, in: Festschrift zur Gedächtnisfeier für Karl Weierstraß 1815–1965 (Wiss. Abh. Arbeitsgemeinsch. f. Forschg. d. Landes Nordrhein-Westf., Bd. 33; herausgeg. von H. Behnke und K. Kopfermann, Westdeutscher Verlag, Köln u. Opladen) pp. 183–240.

Hosotani, Y. (1981), Local scalar fields equivalent to Nambu-Goto strings, Phys. Rev. Lett. 47, 399–401.

Itzykson, C. and J.-B. Zuber (1980), Quantum Field Theory (McGraw-Hill Book, Co., N.Y.).

Jackiw, R. (1980), Quantum meaning of classical field theory, Rev. Mod. Phys. 49, 681–706.

Juvet, G. (1937), Mécanique analytique et mécanique ondulatoire (Mém. Sci. Math. fasc. 83; Gauthier-Villars, Paris).

Kafiev, Yu.Y. (1981), Four-dimensional σ-models and geometry of Yang–Mills fields, Nucl. Phys. B178, 177–196.

Kähler, E. (1934), Einführung in die Theorie der Systeme von Differentialgleichungen (Hamburger Mathem. Einzelschriften 16; reprinted by Chelsea Publ. Comp., N.Y. 1949).

Kalb, M. and P. Ramond (1974), Classical direct interstring action, Phys. Rev. 9D, 2273–2284.

Kastrup, H.A. (1977), Carathéodory's canonical theory for fields and the problem of quark confinement, Phys. Lett. 70B, 195–200.

Kastrup, H.A. (1978a), An electromagnetic model for charge confinement, Phys. Lett. 77B, 203–207.

Kastrup, H.A. (1978b), On the relation of Regge-trajectories to degenerate classical solutions of equations of motion, Phys. Lett. 78B, 433–437.

Kastrup, H.A. (1979), Relativistic strings and electromagnetic flux tubes, Phys. Lett. 82B, 237–238.

Kastrup, H.A. (1981), Vortices in ferromagnets as an example for caustics in field theories and as a model for quark confinement, in: Statistical Mechanics of Quarks and Hadrons, ed. H. Satz (North-Holland Publ. Co., Amsterdam, N.Y., Oxford) pp. 433–456.

Kastrup, H.A. and M. Rinke (1981), Hamilton–Jacobi theories for strings, Phys. Lett. 105B, 191–196.

Keller, J.B. (1958), Corrected Bohr–Sommerfeld quantum conditions for nonseparable systems, Ann. Physics 4, 180–188.

Kijowski, J. (1972), A construction of algebra of the observables in the classical field theory, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 20, 597–602.

Kijowski, J. (1973), A finite-dimensional canonical formalism in the classical field theory, Commun. Math. Phys. 30, 99–128.

Kijowski, J. (1974), Multiphase spaces and gauge in the calculus of variations, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. Phys. 22, 1219–1225.

Kijowski, J. (1977), Symplectic geometry and second quantization, Reports Math. Phys. 11, 97–109.

Kijowski, J. and W. Szczyrba (1976), A canonical structure for classical field theories, Commun. Math. Phys. 46, 183–206.

Kijowski, J. and W.M. Tulczyjew (1979), A Symplectic Framework for Field Theories (Lect. Notes in Physics 107; Springer-Verlag, Berlin, Heidelberg, N.Y).

Kittel, C. (1971), Introduction to Solid State Physics, 4th ed. (J. Wiley and Sons, N.Y.).

Klötzler, R. (1961), Die Konstruktion Geodätischer Felder in der Variationsrechnung mehrfacher Integrale, S.-B. Sächs. Akad. Wiss. Leipzig Math.-Naturw. Kl., Bd. 104, Heft 6.

Klötzler, R. (1970), Mehrdimensionale Variationsrechnung (Birkhäuser, Basel).

Kneser, H. (1921), Untersuchungen zur Quantentheorie, Math. Ann. 84, 277–302.

Kneser, H. (1940), Homogene Funktionen auf der Grassmannschen Mannigfaltigkeit, Bull. Soc. Math. Gréce 20, 101–103.

Kogut, J.B. (1979), An introduction to lattice gauge theory and spin systems, Rev. Mod. Phys. 51, 659–713.

Kostant, B. (1977), Graded manifolds, graded Lie theory and prequantization, in: Differential Geometrical Methods in Mathematical Physics, Bonn 1975, eds. K. Bleuler and A. Reetz (Lect. Notes in Math. 570; Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 177–306.

Kosterlitz, J.M. and D.J. Thouless (1973), Ordering, metastability and phase transitions in two-dimensional systems, J. Phys. C6, 1181–1203.

Kosterlitz, J.M. (1974), The critical properties of the two-dimensional xy model, J. Phys. C7, 1046–1060.

Kupershmidt, B.A. (1980), Geometry of jet bundles and the structure of Lagrangian and Hamiltonian systems, in: Geometric Methods in Mathematical Physics, eds. G. Kaiser and J. Marsden (Lect. Notes in Math. 775; Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 162–218.

Kuranishi, M. (1957), On E. Cartan's prolongation theorem of exterior differential systems, Amer. J. Math. 79, 1–47.

Kuranishi, M. (1967), Lectures on Involutive Systems of Partial Differential Equations (Publicações da Sociedade de Matemática de São Paulo, São Paulo).

Kyriakopoulos, E. (1982), Real solutions of the Yang–Mills field equations in Minkowski space, J. Math. Phys. 23, 2467–2471.

Landau, L.D. (1937), On the theory of phase transitions, in: Collected Papers of L.D. Landau, ed. D. ter Haar (Pergamon Press, Oxford, 1965) pp. 193–216.

Landau, L.D. and E.M. Lifshitz (1962), The Classical Theory of Fields, 2nd rev. ed. (Course of Theoretical Physics, vol. 2; Pergamon Press, Oxford).

Langer, J.S. (1965), A modified spherical model of a first-order phase transition, Phys. Rev. 137A, 1531–1547.

Lawson Jr., H.B. (1977), The Quantitative Theory of Foliations (Regional Conf. Ser. Math. No. 27; Am. Math. Soc., Providence, R.I.).

Lepage. Th.H.J. (1936a, b), Sur les champs géodésiques du calcul des variations, Acad. Roy. Belg. Bull. Cl. Sci. (5e Sér) 22, 716–729; 1036–1046.

Lepage, Th.H.J. (1941), Sur les champs géodésiques des intégrales multiples, Acad. Roy. Belg. Bull. Cl. Sci. (5e Sér.) 27, 27–46.

Lepage, Th.H.J. (1942a, b), Champs stationnaires, champs géodésiques et formes intégrables, Acad. Roy. Belg. Bull. Cl. Sci. (5e Sér.) 28, 73–92; 247–265.

Lie, S. (1896), Die infinitesimalen Berührungstransformationen der Optik, in: S. Lie, Gesamm. Abh., Bd. 6, eds. F. Engel and P. Heegaard (Teubner, Leipzig; Aschenhoug, Oslo, 1927) pp. 615–617.

Lie, S. and G. Scheffers (1896), Geometrie der Berührungstransformationen, Bd. 1 (B.G. Teubner, Leipzig).

Liesen, A. (1967), Feldtheorie in der Variationsrechnung mehrfacher Integrale, Math. Ann. 171, 194–218; 273–292.
Lund, F. and T. Regge (1976), Unified approach to strings and vortices with soliton solutions, Phys. Rev. D14, 1524–1535.
Mack, G. and V.B. Petkova (1979), Comparison of lattice gauge theories with gauge groups Z_2 and SU(2), Ann. Phys. 123, 442–467.
Mack, G. and V.B. Petkova (1980), Sufficient condition for confinement of static quarks by a vortex condensation mechanism, Ann. Phys. 125, 117–134.
Magg, M. (1982), Die klassische nichtabelsche Eichtheorie im Minkowski Raum, Habilitationsschr. (RWTH Aachen), to appear in Fortschr. d. Physik.
Maison, D. (1979), Some facts about classical non-linear σ-models, Preprint MPI-PAE/PTh 52/79 (Max-Planck-Institut Physik, München).
Mandelstam, S. (1976), Vortices and quark confinement in non-abelian gauge theories, Phys. Reports 23, 245–249.
Maslov, V.P. (1972), Théorie des Perturbations et Méthodes Asymptotiques (Dunod, Paris).
Maslov, V.P. and M.V. Fedoriuk (1981), Semi-Classical Approximations in Quantum Mechanics (Reidel Publ. Co., Dordrecht, Boston, London).
Matsushima, Y. (1972), Differentiable Manifolds (Marcel Dekker, N.Y.).
Mayer, A. (1904, 1906), Über den Hilbertschen Unabhängigkeitssatz in der Theorie des Maximums und Minimums der einfachen Integrale, Math. Ann. 58, 235–248; 62, 335–350.
Misner, Ch.W., K.S. Thorne and J.A. Wheeler (1973), Gravitation (W.H. Freeman, San Francisco).
Moore, M.A. (1972), Renormalization and phase transitions, Lett. Nuovo Cim. 3, 275–280.
Morgenstern, M. (1982), Diplom thesis, Inst. f. Theoretische Physik, RWTH Aachen.
Nambu, Y. (1970), Quark model and the factorization of the Veneziano amplitude, in: Symmetries and Quark Models, ed. R. Chand (Gordon and Breach, N.Y.) pp. 269–278.
Nambu, Y. (1974), Strings, monopoles and gauge fields, Phys. Rev. D10, 4262–4268.
Nambu, Y. (1977), Strings, vortices and gauge fields, in: Quark Confinement and Field Theory, eds. D.R. Stump and D.H. Weingarten (J. Wiley and Sons, N.Y.) pp. 1–12.
Nambu, Y. (1980), Hamilton–Jacobi formalism for strings, Phys. Lett. 92B, 327–330.
Nielsen, H.B. (1970), Contrib. to the XVth Intern. Conf. on High Energy Physics, Kiev 1970, Proc. ed. by V. Shelest et al. (Naukova Dumka Publ., Kiev 1972) p. 723.
Nielsen, H.B. and P. Olesen (1973), Local field theory of the dual string, Nucl. Phys. B57, 367–380.
Noether, E. (1918), Invariante Variationsprobleme, Nachr. d. Königl. Gesellsch. d. Wiss. Göttingen, Math.-physik. Kl., 235–257.
Nôno, T. and F. Mimura (1972), A differential geometric formalism for multiple integral variational problems, Tensor, N. S. 26, 79–85.
Nôno, T. and F. Mimura (1975), Null class of the generalized system in mechanics, Tensor, N. S. 29, 69–81.
Nüsser, A. (1982), Diplom thesis, Inst. f. Theor. Physik, RWTH Aachen.
Ogielski, A.T. (1980), Comments on the "planar-time" dynamics of strings, Phys. Rev. D22, 2407–2411.
Okun, L.B. (1981), Particle physics prospects: August '81, in: Proc. 1981 Intern. Symp. on Lepton and Photon Interactions at High Energies, ed. W. Pfeil (Physik. Inst., Univ. Bonn).
Pars, L.A. (1965), A Treatise on Analytical Dynamics (Heineman, London).
Patashinskij, A.Z. and V.L. Pokrovskij (1964), Second order phase transitions in a Bose fluid, Sov. Phys. JETP 19, 677–691.
Pham Mau Quan (1969), Introduction a la Géométrie des Variétés Différentiables (Dunod, Paris).
Pohlmeyer, K. (1976), Integrable Hamiltonian systems and interactions through quadratic constraints, Commun. Math. Phys. 46, 207–221.
Poston, T. and I. Stewart (1978), Catastrophe Theory and its Applications (Pitman, London).
Prange, G. (1915), Die Hamilton–Jacobische Theorie für Doppelintegrale, Dissert., Göttingen.
Quigg, C. and J.L. Rosner (1979), Quantum mechanics with applications to quarkonium, Phys. Reports 56, 167–235.
Radò, T. (1933), On the problem of Plateau, Ergebn. Math. Grenzgeb. 2, 1–109.
Rebbi, C. (1974), Dual models and relativistic quantum strings, Phys. Reports 12, 1–73.
Rinke, M. (1980), The relation between relativistic strings and Maxwell fields of rank 2, Commun. Math. Phys. 73, 265–271.
Rinke, M. (1981), private communication, to be published.
Rund, H. (1973), The Hamilton–Jacobi theory in the calculus of variations (R.E. Krieger Publ. Co., Huntington, N.Y.).
Salam, A. and J. Strathdee (1978), Supersymmetry and superfields, Fortschr. d. Physik 26, 57–142.
Satake, I. (1975), Linear Algebra (Marcel Dekker, Inc., N.Y.).
Scherk, J. (1975), An introduction to the theory of dual models and strings, Rev. Mod. Phys. 47, 123–164.
Schild, A. (1977), Classical null strings, Phys. Rev. D16, 1722–1726.
Schwartz. A.S. (1979), Instantons and fermions in the field of instanton, Commun. Math. Phys. 64, 233–268.
Schwinger, J. (1962), Gauge invariance and mass. II, Phys. Rev. 128, 2425–2429.
Schwinger, J. (1963), Gauge theories of vector particles, in: Theoretical Physics (Intern. Atom. Energy Agency, Vienna) pp. 89–134.
Shifman, M.A., A.I. Vainshtein and V.I. Zakharov (1979), QCD and resonance physics (I, II), Nucl. Phys. B147, 385–534.
Simon, B. and A.D. Sokal (1981), Rigorous entropy-energy arguments, J. Stat. Phys. 25, 679–694.
Singer, I.M. and J.A. Thorpe (1976), Lecture Notes on Elementary Topology and Geometry (Springer-Verlag, N.Y., Heidelberg, Berlin).
Ślebodziński, W. (1970), Exterior Forms and their Applications (Polish Scient. Publishers, Warszawa).
Smale, S. (1970), Topology and mechanics. I, Invent. math. 10, 305–331.
Sniatycki, J. (1970), On the geometric structure of classical field theory in Lagrangian formulation, Proc. Camb. Phil. Soc. 68, 475–484.
Souriau, J.M. (1970), Structure des Système Dynamique (Dunod, Paris).
Sternberg, S. (1964), Lectures on Differential Geometry (Prentice-Hall, Inc., Englewood Cliffs).

Sternberg, S. (1977), Some recent results on supersymmetry, in: Differential Geometrical Methods in Mathematical Physics, Bonn 1975, eds. K. Bleuler and A. Reetz (Lect. Notes in Math. 570; Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 145–176.
Strocchi, F. and A.S. Wightman (1974), Proof of the charge superselection rule in local relativistic quantum field theory, J. Math. Physics 15, 2198–2224.
Susskind, L. (1970), Dual-symmetric theory of hadrons-I, Nuovo Cim. 69A, 457–495.
Szczyrba, W. (1974), A Lagrangian and canonical formalism in the hydrodynamics of an incompressible fluid, Reports Math. Phys. 6, 289–302.
Szczyrba, W. (1976), Lagrangian formalism in the classical field theory, Ann. Polon. Math. 32, 145–185.
Szczyrba, W. (1977), On the geometrie structure of the set of solutions of Einstein equations, Dissert. Mathem. 150 (Polish Acad. Sciences, Warsaw).
Szczyrba, W. (1981), A Hamiltonian structure of the interacting gravitational and matter fields, J. Math. Phys. 22, 1926–1943.
Takens, F. (1977), Symmetries, conservation laws and variational principles, in: Geometry and Topology, eds. J. Palis and M. do Carmo (Lect. Notes in Math. 597; Springer-Verlag, Berlin, Heidelberg, N.Y.) pp. 581–604.
Takens, F. (1979), a global version of the inverse problem of the calculus of variations, J. Diff. Geometry 14, 543–562.
Takhtadzhyan L.A. and L.D. Faddeev (1974), Essentially nonlinear one-dimensional model of classical field theory, Theor. and Math. Phys. 21, 1046–1057.
Thacker, H.B. (1981), Exact integrability in quantum field theory and statistical systems, Rev. Mod. Phys. 53, 253–285.
Thirring, W. (1978), Classical Dynamical Systems (Springer, N.Y.).
't Hooft, G. (1976a), Symmetry breaking through Bell–Jackiw anomalies, Phys. Rev. Lett. 37, 8–11.
't Hooft, G. (1976b), Computation of the quantum effects due to a four-dimensional pseudoparticle, Phys. Rev. D14, 3432–3450.
't Hooft, G. (1976c), Gauge theories with unified weak, electromagnetic and strong interactions, in: High Energy Physics, Proc. EPS Intern. Conf., Palermo 1975, ed. A. Zichichi (Editrice Compositori, Bologna) pp. 1225–1249.
t' Hooft, G. (1978), On the phase transitions towards permanent quark confinement, Nucl. Phys. B138, 1–25.
Trautman, A. (1967), Noether equations and conservation laws, Commun. Math. Phys. 6, 248–261.
Trautman, A. (1972), Invariance of Lagrangian Systems, in: General Relativity, Papers in honor of J.L. Synge, ed. L. O'Raifeartaigh (Clarendon Press, Oxford) pp. 85–99.
Tulczyjew, W.M. (1974), Poisson brackets and canonical manifolds, Bull. Acad. Polon. Sci. Sér. Sci. Math. Astronom. 22, 931–935.
Van Hove, L. (1945a), Sur la construction des champ de DeDonder–Weyl par la méthode des charactéristiques, Acad. Roy. Belg. Bull. Cl. Sci. (5e Sér.) 31, 278–285.
Van Hove, L. (1945b), Sur les champs de Carathéodory et leur construction par la méthode des characteristiques, Acad. Roy. Belg. Bull. Cl. Sci. (5e Sér.) 31, 625–638.
Velte, W. (1953), Zur Variationsrechnung mehrfacher Integrale in Parameterdarstellung, Mitt. Math. Sem. Giessen, Heft 45, 1–33.
Velte, W. (1954), Zur Variationsrechnung mehrfacher Integrale, Math. Zeitschr. 60, 367–383.
Vessiot, E. (1906), Sur l'interprétation mécanique des transformations de contact infinitésimales, Bull. Soc. Math. France 34, 230–269.
Vessiot, E. (1909), Essai sur la propagation par ondes, Ann. Sci. Ecole Norm. Sup. 26, 405–448.
Vessiot, E. (1912), Sur la théorie des multiplicités et le calcul des variations, Bull. Soc. Math. France 40, 68–139.
Vessiot, E. (1913), Sur la propagation par ondes et sur le problème de Mayer, J. Math. Pures Appl. (Sér. 6) 9, 39–76.
Vessiot, E. (1914), Sur les invariants intégraux de la propagation par ondes, Bull. Soc. Math. France 42, 142–167.
Vinogradov, A.M. (1977), On the algebro-geometric foundation of Lagrangian field theory, Sov. Math. Dokl. 18, 1200–1204.
Vinogradov, A.M. (1978), A spectral sequence associated with a nonlinear differential equation, and algebro-geometric foundation of Lagrangian field theory with constraint, Sov. Math. Dokl. 19, 144–148.
Volterra, V. (1890), Sopra una estensione della teoria Jacobi–Hamilton del Calcolo delle variazoni, Atti Accad. Naz. Lincei Rend. (Ser. IV) VI. 127–138.
von Rieth, J. (1982), The Hamilton–Jacobi theory of DeDonder and Weyl applied to some relativistic field theories, to appear in J. Math. Phys.
von Rieth, J. and H.A. Kastrup (1983), Properties of exterior differential systems associated with variational problems in field theories, in preparation.
Voros, A. (1976), Semi-classical approximations, Ann. Inst. Henri Poincaré A 24, 31–90.
Warner, F.W. (1971), Foundations of Differentiable Manifolds and Lie Groups (Scott, Foresman and Co., Glenview, London).
Wegner, F.J. (1976), The critical state, general aspects, in: Phase Transitions and Critical Phenomena, vol. 6, eds. C. Domb and M.S. Green (Academic Press, London, N.Y., San Francisco) pp. 8–124.
Weinstein, A. (1977), Lectures on Symplectic Manifolds (Regional Conf. Series in Math. vol. 29; Am. Math. Soc., Providence).
Weyl, H. (1934), Oberservations on Hilbert's independence theorem and Born's quantization of field equations, Phys. Rev. 46, 505–508.
Weyl, H. (1935), Geodesic fields in the calculus of variation for multiple integrals, Ann. Math. 36, 607–629.
Whittaker, E.T. (1959), A Treatise on the Analytical Dynamics of Particles and Rigid Bodies, 4th ed. (Cambridge Univ. Press, Cambridge).
Wilkins Jr., J.E. (1944), Multiple integral problems in parametric form in the calculus of variations, Ann. Math. 45, 312–334.
Wilson, K.G. (1971), Renormalization group and critical phenomena II. Phase-space cell analysis of critical behavior, Phys. Rev. B4, 3184–3205.
Wilson, K.G. and J. Kogut (1974), The renormalization group and the ε-expansion, Phys. Reports 12, 75–200.
Wolkow, D.M. (1935), Über eine Klasse von Lösungen der Diracschen Gleichung, Zeitschr. f. Physik 94, 250–260.
Wu., T.T. and C.N. Yang (1975), Concept of nonintegrable phase factors and global formulation of gauge fields, Phys. Rev. D12, 3845–3857.
Yoneya, T. (1978), $Z(N)$ topological excitations in Yang–Mills theories: duality and confinement, Nucl. Phys. B144, 195–218.
Zumino, B. (1976), Supersymmetry, in: Gauge Theories and Modern Field Theory, eds. R. Arnowitt and P. Nath (The MIT Press, Cambridge, Mass. and London) pp. 256–280.

References added in proof

As to my remarks in section 5.9 concerning the work of Kupershmidt I have to point out the following papers, which are relevant in the context of problem 6 in ch. 9, too:

Gibbons, J., D.D. Holm and B. Kupershmidt (1983), The Hamilton structure of classical chromohydrodynamics, Physica 6D, 179–194.

Holm, D.D. and B.A. Kupershmidt (1983), Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity, Physica 6D, 347–363.

Holm, D.D. and B.A. Kupershmidt (1983), Noncanonical Hamiltonian formulation of ideal magnetohydrodynamics, Physica 7D, 330–333.

Kupershmidt, B.A. (1983), On dual spaces of differential Lie algebras, Physica 7D, 334–337.

These papers contain further references.

As to the topic 6 of ch. 9 I would like to add the following references:

Faddeev, L.D. (1980), A Hamiltonian interpretation of the inverse scattering method, in: Solitons, eds. R.K. Bullough and P.J. Caudrey (Topics in current physics 17; Springer-Verlag, Heidelberg, Berlin, N.Y.) pp. 339–354.

Glimm, J. and A. Jaffe (1981), Quantum physics, a functional integral point of view (Springer-Verlag, N.Y., Heidelberg, Berlin).

Kodama, Y. and M. Wadati (1976), Theory of canonical transformations for nonlinear evolution equations. I, Progr. Theor. Physics 56, 1740–1755.

Kodama, Y. (1977), Theory of canonical transformations for nonlinear evolution equations. II, Progr. Theor. Physics 57, 1900–1916.

Kodama, Y. (1977), A canonical transformation for equation of the form $\phi_{xt} = h(\phi)$, Journ. Phys. Soc. Japan 42, 1029–1033.

Marsden, J. and A. Weinstein (1983), Coadjoint orbits, vortices and Clebsch variables for incompressible fluids, Physica 7D, 305–323.