Field theory with higher derivatives-Hamiltonian structure

1969 J. Phys. A: Gen. Phys. 2 304

(http://iopscience.iop.org/0022-3689/2/3/009)

J. PHYS. A (GEN. PHYS.), 1969, SER. 2, VOL. 2. PRINTED IN GREAT BRITAIN

Field theory with higher derivatives—Hamiltonian structure

L. M. C. COELHO DE SOUZA and P. R. RODRIGUES†

Departamento de Matemática, Faculdade de Filosofia, Universidade Federal do Rio Grande do Sul, Pôrto Alegre, Brazil

MS. received 26*th August* 1968, *in revised form* 2*nd December* 1968

Abstract. Canonical formalism, including Hamilton's equations, transformation theory and Poisson brackets for classical field theory with higher time and/or spatial derivatives, is established.

1. Introduction

Since the early work of Podolski and co-workers (Podolski 1942, Podolski and Kikuchi 1944, Podolski and Schwed 1948), where an electromagnetic theory with second-order derivatives was introduced, many attempts (Chang 1948, Green 1948, Katayama 1954, Taniuti 1955, 1956) have been made to exploit a theory of fields with higher derivatives, with the clear intention of trying to generalize the theory of Podolski. In these attempts two different alternatives have been followed and, as far as we know, the present situation concerning this problem is as follows.

(i) Hamilton's equations for the special case where the Lagrangian contains derivatives up to second order have been derived by Podolski.

(ii) For a special class of field theory with higher derivatives, canonical equations were derived by Chang, without either an explicit definition of the canonical variables or a development of a transformation theory. The problem is reduced to a usual one where the Lagrangian contains derivatives up to first order and the variation of the action integral is performed under $s-1$ subsidiary conditions (a set of field coordinates are time derivatives of some other coordinates). The unknown multipliers introduced by the process are identified with $s-1$ momenta, which are not directly defined—actually these momenta are solutions of a system of $s-1$ first-order differential equations. That is the reason why (although the canonical equations so obtained are formally identical with ours) the construction of a suitable Hamiltonian (which in general is a hard task) becomes virtually impossible for large s. As one of the main applications of this theory (an approach to the non-local interaction as infinite sums of interactions of a higher derivative ($s \to \infty$)) concerns precisely these large s, it is perhaps explained why this formulation has not been used in subsequent papers dealing with the problem.

(iii) De Wett (1948) followed a different alternative defining canonical variables and Hamiltonian formalism. The following criticisms are valid: (*a*) the Hamiltonian formalism does not contain the simplest one as a particular case (i.e. when the field variables do not depend on the spatial coordinates, the usual formalism of the mechanics of this higher-order derivative is not obtained); (*b*) a transformation which leaves these equations invariant is not available (up to now).

(iv) Canonical transformations and Poisson brackets for such theories have not been studied.

(v) A Hamiltonian formalism was developed by Thielheim (1967). It will be discussed in detail in appendix 2 because of its formal resemblance to ours.

It is the purpose of this paper (i) to establish a Hamiltonian formulation with canonical variables explicitly defined, and (ii) to prove, via a transformation theory, invariance of the formalism and to define Poisson brackets for a classical field theory described by a Lagrangian density

$$\mathscr{L} = \mathscr{L}(\varphi_\alpha, \varphi_{\alpha;i_1}, \ldots, \varphi_{\alpha;i_1\ldots i_s}, x_0, x_1, \ldots, x_n)$$

† Conselho Nacional de Pesquisas Fellow.

where the $\varphi_\alpha = \varphi_\alpha(x_0, x_1, ..., x_n)$ are the field coordinates (with $\alpha = 1, 2, ..., N$), x_0, $x_1, ..., x_n$ are independent coordinates and

$$\varphi_{\alpha;i_1...i_m} = \frac{\partial^m \varphi_\alpha}{\partial x_{i_1} \dots \partial x_{i_m}}.$$

Let us consider the functional

$$J = \int \dots \int_R \mathscr{L}(\varphi_\alpha, \varphi_{\alpha;i_1...}, \varphi_{\alpha;i_1...i_s}, x_0, \dots, x_n)\, \mathrm{d}x_0 \dots \mathrm{d}x_n. \tag{1}$$

It is assumed that the integrand $\mathscr{L}$ has continuous derivatives up to order s with respect to all its arguments.

Assuming that the region R remains fixed, while the functions are varied in such a way as

$$\begin{aligned} \varphi_\alpha(x) &\to \varphi_\alpha^*(x) = \varphi_\alpha(x) + \epsilon f_\alpha(x) + \text{higher powers of } \epsilon \\ \varphi_{\alpha;i_1}(x) &\to \overset{*}{\varphi}_{\alpha;i_1}(x) = \varphi_{\alpha;i_1}(x) + \epsilon f_{\alpha;i_1}(x) + \text{higher powers of } \epsilon \\ &\vdots \qquad\qquad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ \varphi_{\alpha;i_1...i_s}(x) &\to \overset{*}{\varphi}_{\alpha;i_1...i_s}(x) = \varphi_{\alpha;i_1...i_s}(x) + \epsilon f_{\alpha;i_1...i_s}(x) + \text{higher powers of } \epsilon. \end{aligned} \tag{2}$$

(We use $\varphi_\alpha(x)$, $\varphi_{\alpha;i_1}(x)$, $\varphi_{\alpha;i_1...i_s}(x)$, etc., instead of $\varphi_\alpha(x_0, x_1, ..., x_n)$, $\varphi_{\alpha;i_1}(x_0, x_1, ... x_n)$, $\varphi_{\alpha;i_1...i_s}(x_0, x_1, ... x_n)$, etc.)

Let $\delta J = J^* - J$ be the variation of functional (1), corresponding to the transformation (2), i.e.

$$\begin{aligned} \delta J = \int_{R^n} \Big[&\mathscr{L}\{\varphi_\alpha(x) + \epsilon f_\alpha, \varphi_{\alpha;i_1}(x) + \epsilon f_{\alpha;i_1}(x), \dots, \varphi_{\alpha;i_1...i_s}(x) + \epsilon f_{\alpha;i_1...i_s}(x), x\} \\ &- \mathscr{L}\{\varphi_\alpha(x), \varphi_{\alpha;i_1}(x), \dots, \varphi_{\alpha;i_1...i_s}(x), x\} \Big]\, \mathrm{d}^n x. \end{aligned}$$

Using Taylor's theorem, we find the variation of functional (1) to be

$$\delta J = \epsilon \int_{R^n} \left\{ \frac{\partial \mathscr{L}}{\partial \varphi_\alpha} f_\alpha + \sum_0^n{}_{i_1} \left(\frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1}} \right) f_{\alpha;i_1} + \dots + \sum_0^n{}_{i_1...i_s} \left(\frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1...i_s}} \right) f_{\alpha;i_1...i_s} \right\} \mathrm{d}^n x. \tag{3}$$

In order to avoid unnecessary calculations we restrict ourselves to the case where R^n is a hyperparallelepiped. (The result obtained by the repeated use of Green's theorem will be the same whatever the shape of R^n.) Integrating by parts and rearranging equation (3) we obtain

$$\begin{aligned} \delta J = \epsilon \int_{R^n} &\left[\left\{ \frac{\partial \mathscr{L}}{\partial \varphi_\alpha} - \sum_0^n{}_{i_1} \frac{\mathrm{d}}{\mathrm{d}x_{i_1}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1}} + \dots + (-1)^s \sum_0^n{}_{i_1...i_s} \frac{\mathrm{d}^s}{\mathrm{d}x^s_{i_1...i_s}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1...i_s}} \right\} f_\alpha \right] \mathrm{d}^n x \\ &+ \epsilon \left[\left\{ \sum_0^n{}_{i_1} \int_{R^{n-1}} \left(\frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1}} f_\alpha \right) \mathrm{d}^{n-1} x \right\} \right. \\ &+ \left\{ \sum_0^n{}_{i_1 i_2} \int_{R^{n-2}} \left(\frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1 i_2}} f_{\alpha;i_1} - \frac{\mathrm{d}}{\mathrm{d}x_{i_1 i_2}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1 i_2}} f_\alpha \right) \mathrm{d}^{n-2} x \right\} \\ &+ \dots \\ &+ \left\{ \sum_0^n{}_{i_1...i_s} \left(\frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1...i_s}} f_{\alpha;i_1...i_{s-1}} - \frac{\mathrm{d}}{\mathrm{d}x_{i_1...i_s}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1...i}} f_{\alpha;i_1...i_{s-2}} \right. \right. \\ &\left. \left. \left. + \dots + (-1)^{s-1} \frac{\mathrm{d}^{s-1}}{\mathrm{d}x^{s-1}_{..1i.i_s}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1...i_s}} f_\alpha \right) \right\} \right] \end{aligned}$$

where $d^{n-1}x$ is the volume element in the space R^{n-1} defined by $R^{n-1}U\{x_{i_1}\} = R^n$, $d^{n-2}x$ is the element in R^{n-2} defined by $R^{n-2}U\{x_{i_1}\}U\{x_{i_2}\} = R^n$ etc. The curly bracket means, in this case, that the function is calculated in the boundary of R^n.

We obtain the Euler–Lagrange equations

$$\frac{\partial \mathscr{L}}{\partial \varphi_\alpha} - \sum_0^n {}_{i_1} \frac{d}{dx_{i_1}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1}} + \sum_0^n {}_{i_1,i_2} \frac{d}{dx_{i_1}} \frac{d}{dx_{i_2}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1 i_2}}$$

$$- \dots (-1)^s \sum_0^n {}_{i_1 \dots i_s} \frac{d}{dx_{i_1}} \dots \frac{d}{dx_{i_s}} \frac{\partial \mathscr{L}}{\partial \varphi_{\alpha;i_1 \dots i_s}} = 0 \tag{4}$$

provided

$$f_\alpha = f_{\alpha;i_1} = \dots = f_{\alpha;i_1 \dots i_{s-1}} = 0$$

in the hypersurface bounding R^n, or equivalently

$$\delta\varphi_{\alpha;i_1 \dots i_m} = 0, \qquad m = 1, \dots, s-1.$$

2. Hamilton's equations

In order to obtain a Hamiltonian formulation, let us single out one variable (x_0) which may be called the time. It will be useful to introduce the following generalization of functional (or variational) derivatives. If

$$\mathscr{F} = \mathscr{F}(\varphi_\alpha, \varphi_{\alpha;i_1}, \dots, \varphi_{\alpha;\, i_1 \dots i_s}, x_0, \dots x_n)$$

is a function of the field coordinates, its derivatives and the independent variables x_0, $x_1, \dots, x_n$, then the s-functional derivative of the functional

$$F = \int \dots \int \mathscr{F} \, dx_1 \dots dx_n \equiv \int \mathscr{F} \, dv$$

with respect to φ_α is denoted by $\Delta_s F/\Delta_s \varphi_\alpha$ and by definition is

$$\frac{\Delta_s F}{\Delta_s \varphi_\alpha} \overset{\text{def}}{=} \frac{\partial \mathscr{F}}{\partial \varphi_\alpha} - \sum_1^n {}_{i_1} \frac{d}{dx_{i_1}} \frac{\partial \mathscr{F}}{\partial \varphi_{\alpha;i_1}} + \dots + (-1)^s \sum_1^n {}_{i_1 \dots i_s} \frac{d}{dx_{i_1}} \dots \frac{d}{dx_{i_s}} \frac{\partial \mathscr{F}}{\partial \varphi_{\alpha;i_1 \dots i_s}}. \tag{5}$$

(In the case where $s = 1$, we have the correct usual functional derivative.) We shall omit the s subscript of $\Delta_s F/\Delta_s \varphi_\alpha$ whenever possible without confusion.

With the aid of this definition, equation (4) may be written as

$$\sum_0^s {}_l (-1)^l \frac{d^l}{dx_0{}^l} \frac{\Delta L}{\Delta \varphi_\alpha{}^{(l)}} = 0 \tag{4'}$$

where

$$L = \int \dots \int \mathscr{L} \, dx_1 \dots dx_n$$

and

$$\varphi_\alpha{}^{(l)} = \frac{d^l \varphi_\alpha}{dx_0{}^l}.$$

Equation (4′) shows a very close formal resemblance to the equation

$$\sum_0^s {}_l (-1)^l \frac{d^l}{dt^l} \frac{\partial L}{\partial q_\alpha{}^{(l)}} = 0$$

which, as is well known, describes a (generalized) system of mechanics including higher time derivatives of the coordinates $q_\alpha(t)$. By analogy, we define the conjugate momenta as

$$\pi_{\alpha/m} \overset{\text{def}}{=} \sum_0^{s-m} {}_l (-1)^l \frac{d^l}{dx_0{}^l} \frac{\Delta L}{\Delta \varphi_\alpha{}^{(l+m)}} \tag{6}$$

i.e. π_α, where $\alpha = 1, 2, \ldots, s$, is a function of φ and its time derivative up to order $2s-\alpha$:

$$\pi_\alpha = \pi_\alpha\left(\varphi, \frac{\partial\varphi}{\partial x_0}, \frac{\partial^2\varphi}{\partial {x_0}^2}, \ldots, \frac{\partial^{2s-\alpha}\varphi}{\partial {x_0}^{2s-\alpha}}\right).$$

Therefore, the sets of coordinates $\gamma_s = \{\varphi, \ldots, \varphi^{(s-1)}\}$ and $\theta_s = \{\pi_1, \ldots, \pi_s\}$ are independent, in general. Even if they are not, when s is the highest order of derivation, there will always exist an $s' < s$, such that the γ_s and θ_s will be independent in a theory where the highest order of derivation is s'. The conjugate momenta so defined are solutions of the system proposed by Chang (1948).

We are now able to define the Hamiltonian H:

$$H = \int \ldots \int \varkappa \, \mathrm{d}x_1 \ldots \mathrm{d}x_n \overset{\mathrm{def}}{=} \int \ldots \int \left(\sum_\alpha \sum_{1}^{s}{}_m \, \pi_{\alpha/m}\varphi_\alpha^{(m)} - \mathscr{L}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n. \tag{7}$$

From equation (7) and using definition (5) we obtain

$$\frac{\Delta H}{\Delta\varphi_\alpha^{(m)}} = \pi_{\alpha/m} - \frac{\Delta L}{\Delta\varphi_\alpha^{(m)}}$$

which according to equation (6) can be written

$$\begin{aligned}\frac{\Delta H}{\Delta\varphi_\alpha^{(m)}} &= \sum_{0}^{s-m}{}_l \, (-1)^l \frac{\mathrm{d}^l}{\mathrm{d}{x_0}^l}\frac{\Delta L}{\Delta\varphi_\alpha^{(l+m)}} - \frac{\Delta L}{\Delta\varphi_\alpha^{(m)}} \\ &= \sum_{1}^{s-m}{}_l \, (-1)^l \frac{\mathrm{d}^l}{\mathrm{d}{x_0}^l}\frac{\Delta L}{\Delta\varphi_\alpha^{(l+m)}} \\ &= -\frac{\mathrm{d}}{\mathrm{d}x_0}\sum_{0}^{s-(m+1)}{}_l \, (-1)^l \frac{\mathrm{d}^l}{\mathrm{d}{x_0}^l}\frac{\Delta L}{\Delta\varphi_\alpha^{(l+m-1)}}.\end{aligned}$$

Finally, again using definition (6), we have

$$\frac{\Delta H}{\Delta\varphi_\alpha^{(m)}} = -\frac{\mathrm{d}}{\mathrm{d}x_0}\pi_{\alpha/m+1} \tag{8}$$

for $m = 0, 1, \ldots, s-1$ (since $\Delta H/\Delta\varphi_\alpha^{(s)}$ is zero because of (6)). Finally

$$\frac{\Delta H}{\Delta\pi_{\alpha/m+1}} = \varphi_\alpha^{(m+1)} = \frac{\mathrm{d}}{\mathrm{d}x_0}\varphi_\alpha^{(m)} \tag{9}$$

for $m = 0, 1, \ldots, s-1$ since H is independent of $\pi_{\alpha/0}$, which can be proved to be trivially zero, because of (4′). It is obvious that

$$\frac{\partial\mathscr{H}}{\partial x_\nu} = -\frac{\partial\mathscr{L}}{\partial x_\nu}.$$

3. Canonical transformations—Poisson brackets

Let us consider a (time-independent) transformation of the type

$$\varphi_\alpha^{(m)} = \varphi_\alpha^{(m)}[\Phi_\beta, \Phi_\beta^{(1)}, \ldots, \Phi_\beta^{(s-1)}, \Pi_{\beta/1}, \ldots, \Pi_{\beta/s}] \tag{10}$$

$$\pi_{\alpha/m} = \pi_{\alpha/m}[\Phi_\beta, \Phi_\beta^{(1)}, \ldots, \Phi_\beta^{(s-1)}, \Pi_{\beta/1}, \ldots, \Pi_{\beta/s}] \tag{10′}$$

to lead us from the set $\{\varphi, \pi\}$ to another set of coordinates $\{\Phi, \Pi\}$ and such that, provided equations (8) and (9) are satisfied, we obtain

$$-\frac{\mathrm{d}}{\mathrm{d}x_0}\Pi_{\beta/m+1} = \frac{\Delta K}{\Delta\varphi_\beta^{(m)}} \tag{8′}$$

$$\frac{\mathrm{d}}{\mathrm{d}x_0}\Phi_\beta^{(m)} = \frac{\Delta K}{\Delta\Pi_{\beta/m+1}} \tag{9′}$$

where

$$K = [\Pi, \Phi] = H[\pi[\Pi, \Phi], \varphi[\Pi, \Phi]]$$

and $m = 0, 1, \ldots, s-1$; $\beta = 1, 2, \ldots, N$.

If we suppose transformations (10) and (10′) already performed and expand (see appendix 1) equations (8) and (9), we have

$$\sum_1^N{}_\beta \sum_0^{s-1}{}_r \int \ldots \int \left(\frac{\Delta K}{\Delta\Phi_\beta^{(r)}} \frac{\Delta\Phi_\beta^{(r)}}{\Delta\varphi_\alpha^{(m)}} + \frac{\Delta K}{\Delta\Pi_{\beta/r+1}} \frac{\Delta\Pi_{\beta/r+1}}{\Delta\varphi_\alpha^{(m)}}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$
$$= -\sum_1^N{}_\beta \sum_0^{s-1}{}_r \int \ldots \int \left(\frac{\Delta\pi_{\alpha/m+1}}{\Delta\Phi_\beta^{(r)}} \frac{\mathrm{d}\Phi_\beta^{(r)}}{\mathrm{d}x_0} + \frac{\Delta\pi}{\Delta\Pi_{\beta/r+1}} \frac{\mathrm{d}\Pi_{\beta/r+1}}{\mathrm{d}x_0}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n \qquad (8'')$$

$$\sum_1^N{}_\beta \sum_0^{s-1}{}_r \int \ldots \int \left(\frac{\Delta K}{\Delta\Phi_\beta^{(r)}} \frac{\Delta\Phi_\beta^{(r)}}{\Delta\pi_{\alpha/m+1}} + \frac{\Delta K}{\Delta\Pi_{\beta/r+1}} \frac{\Delta\Pi_{\beta/r+1}}{\Delta\pi_{\alpha/m+1}}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n$$
$$= \sum_1^N{}_\beta \sum_0^{s-1}{}_r \int \ldots \int \left(\frac{\Delta\varphi_\alpha^{(m)}}{\Delta\Phi_\beta^{(r)}} \frac{\mathrm{d}\Phi_\beta^{(r)}}{\mathrm{d}x_0} + \frac{\Delta\varphi_\alpha^{(m)}}{\Delta\Pi_{\beta/r+1}} \frac{\mathrm{d}\Pi_{\beta/r+1}}{\mathrm{d}x_0}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n. \qquad (9'')$$

Substituting (8′) and (9′) into (8″) and (9″) and then equating the coefficients of the field variables Φ and Π, we obtain the canonical conditions

$$\frac{\Delta\Phi_\beta^{(r)}}{\Delta\varphi_\alpha^{(m)}} = \frac{\Delta\pi_{\alpha/m+1}}{\Delta\Pi_{\beta/r+1}}; \qquad \frac{\Delta\Phi_\beta^{(r)}}{\Delta\pi_{\alpha/m+1}} = -\frac{\Delta\varphi_\alpha^{(m)}}{\Delta\Pi_{\beta/r+1}};$$
$$\frac{\Delta\Pi_{\beta/r+1}}{\Delta\varphi_\alpha^{(m)}} = -\frac{\Delta\pi_{\alpha/m+1}}{\Delta\Phi_\beta^{(r)}}; \qquad \frac{\Delta\Pi_{\beta/r+1}}{\Delta\pi_{\alpha/m+1}} = \frac{\Delta\varphi_\alpha^{(m)}}{\Delta\Phi_\beta^{(r)}}. \qquad (11)$$

It is easy (although fastidious) to prove that transformation (11) leaves invariant the form

$$[A, B] = \sum_1^N{}_\alpha \sum_0^{s-1}{}_m \int \ldots \int \left(\frac{\Delta A}{\Delta\varphi_\alpha^{(m)}} \frac{\Delta B}{\Delta\pi_{\alpha/m+1}} - \frac{\Delta A}{\Delta\pi_{\alpha/m+1}} \frac{\Delta B}{\Delta\varphi_\alpha^{(m)}}\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n \qquad (12)$$

if A and B are functionals of the field variables. Besides,

$$\frac{\mathrm{d}A}{\mathrm{d}t} = \frac{\partial A}{\partial t} + [A, H]$$

which allows us to look at (12) as a definition of the Poisson bracket of A and B. Some of its main properties are

$$[\varphi_\alpha^{(m)}, \varphi_\beta^{(r)}] = [\pi_{\alpha/m}, \pi_{\beta/r}] = 0$$
$$[\varphi_\alpha^{(m)}(x_0, x_1, \ldots x_n), \pi_{\beta/r}(x_0, x_1', \ldots, x_n')] = \delta_{\alpha\beta}\delta_{m,r-1}\delta(x_1 - x_1') \ldots \delta(x_n - x_n')$$
$$\frac{\mathrm{d}}{\mathrm{d}x_0}\varphi_\alpha^{(m)} = [\varphi_\alpha^{(m)}, H]$$
$$-\frac{\mathrm{d}\pi_{\alpha/m+1}}{\mathrm{d}x_0} = [H, \pi_{\alpha/m+1}]$$

which maintain the analogy with the usual theory.

Appendix 1

Let us suppose that $\mathscr{F}$ is a function of φ, $\varphi^{(1)}$, ..., $\varphi^{(s)}$ and its spatial derivatives and $F = \int \mathscr{F}\, \mathrm{d}v$. If we perform a variation $\Delta\varphi^{(l)}$, then we have

$$\Delta F = \sum_l \int \left(\frac{\partial\mathscr{F}}{\partial\varphi^{(l)}} \Delta\varphi^{(l)} + \sum_i \frac{\partial\mathscr{F}}{\partial\varphi_i^{(l)}} \Delta\varphi_i^{(l)} + \ldots\right) \mathrm{d}v.$$

After k successive partial integrations on the $(k+1)$th term, this expression becomes

$$\Delta F = \sum_l \int \left(\frac{\Delta F}{\Delta\varphi^{(l)}}\right) (\Delta\varphi^{(l)}(x))\, dv. \tag{A1}$$

Now, let us uppose that the $\varphi^{(l)}$ are functionals of ψ, $\psi^{(1)}$, ..., $\psi^{(s)}$ and its spatial derivatives. If we perform a $\Delta\psi^{(l)}$ variation on $\psi^{(l)}$, we have

$$(\Delta\varphi^{(l)})x = \sum_r \int \left(\frac{\Delta\varphi^l(x)}{\Delta\psi^{(r)}(x')}\right) (\Delta\psi^{(r)}(x'))\, dv'. \tag{A2}$$

But, if F is also a functional of ψ,

$$\Delta F = \sum_r \int \frac{\Delta F}{\Delta\psi^{(r)}} \Delta\psi^{(r)}\, dv. \tag{A3}$$

Substituting (A2) into (A1) and comparing the expression obtained with (A3), we obtain the 'chain rule' for our generalized functional

$$\frac{\Delta F}{\Delta\psi^{(r)}} = \sum_l \int \frac{\Delta F}{\Delta\varphi^{(l)}} \frac{\Delta\varphi^{(l)}}{\Delta\psi^{(r)}}\, dv$$

which is used in expanding equations (8) and (9).

Appendix 2

A Hamiltonian formalism was presented by Thielheim (1967), but, although his Hamilton equations are quite similar to ours, there is a fundamental difference between the two definitions of the variable π. Using our symbols, his definition is written

$$\pi_{\alpha/m} = \sum_{0}^{n}{}_j \sum_{0}{}_k (-1)^{j+k} \binom{\beta+j+k}{j} \frac{\partial}{\partial x_{i_1}} \cdots \frac{\partial}{\partial x_{i_j}} \left[\frac{\partial^k}{\partial x_0{}^k} \left(\frac{\partial \mathscr{L}}{\partial \varphi^{(2m+k-1)}_{\alpha;i_1\ldots i_j}}\right)\right].$$

This definition can be said to be a linear combination of spatial derivatives of

$$\frac{\partial \mathscr{L}}{\partial \varphi^{(2m+k-1)}_{\alpha;i_1\ldots i_j}}$$

which can be written as

$$\pi_{\alpha/m} = \text{linear combination of spatial derivatives of} \frac{\partial \mathscr{L}}{\partial \varphi^{(2m-1)}_{\alpha;i_1\ldots i_j}}$$

$$+ \text{linear combination of spatial derivatives of} \frac{\mathrm{d}}{\mathrm{d}x_0} \frac{\partial \mathscr{L}}{\partial \varphi^{(2m)}_{\alpha;i_1\ldots i_r}}$$

$$+ \ldots .$$

In the same manner, if we write our definition of the π as

$$\pi_{\alpha/m} = \sum_{0}^{s-m}{}_l (-1)^l \frac{\mathrm{d}^l}{\mathrm{d}x_0} \left(\frac{\partial \mathscr{L}}{\partial \varphi^{(l+m)}_{\alpha}} - \sum_{1}^{n}{}_{i_1} \frac{\mathrm{d}}{\mathrm{d}x_{i_1}} \frac{\partial \mathscr{L}}{\partial \varphi^{(l+m)}_{\alpha;i_1}} + \ldots \right.$$

$$\left. +(-1)^s \sum_{1}^{n}{}_{i_1\ldots i_s} \frac{\mathrm{d}}{\mathrm{d}x_{i_1}} \cdots \frac{\mathrm{d}}{\mathrm{d}x_{i_s}} \frac{\partial \mathscr{L}}{\partial \varphi^{(l+m)}_{\alpha;i_1\ldots i_s}}\right)$$

we can also say that our $\pi_{\alpha/m}$ are a linear combination of the spatial derivatives of

$$\frac{\partial \mathscr{L}}{\partial \varphi^{(m)}_{\alpha;i_1\ldots i_r}} + \ldots +$$

a linear combination of the spatial derivatives of

$$\frac{\mathrm{d}^{s-m}}{\mathrm{d}x_0^{s-m}} \frac{\partial \mathscr{L}}{\partial \varphi^{(s)}_{\alpha; i_1 \dots i_r}}.$$

We are now able to compare the two definitions, and it is clear that they are equivalent only if $m = 2m-1$ (or $m = 1$), that is, the usual case.

Besides this difference, which leads to two fundamentally different results, the work in question does not develop a transformation theory and the author himself calls attention to the fact that his transformed equations of motion can not in general be reduced to the form of the Lagrange equation.

References

Chang, T. S., 1946, *Proc. Camb. Phil. Soc.*, **42**, 132–8.
—— 1948, *Proc. Camb. Phil. Soc.*, **44**, 76–86.
Green, A. E. S., 1948, *Phys. Rev.*, **73**, 26–9.
Katayama, Y., 1954, *Prog. Theor. Phys., Japan*, **10**, 31–56.
Podolski, B., 1942, *Phys. Rev.*, **62**, 68–71.
Podolski, B., and Kikuchi, 1944, *Phys. Rev.*, **65**, 228–35.
Podolski, B., and Schwed, P., 1948, *Rev. Mod. Phys.*, **20**, 40–50.
Taniuti, T., 1955, *Prog. Theor. Phys., Japan*, **13**, 505–21.
—— 1956, *Prog. Theor. Phys., Japan*, **15**, 19–36.
Thielheim, K. O., 1967, *Proc. Phys. Soc.*, **91**, 798–803.
de Wett, J. S., 1948, *Proc. Camb. Phil. Soc.*, **44**, 546–59.