Math. Proc. Camb. Phil. Soc. (1981), **90**, 537
Printed in Great Britain

The Variational Principle and natural transformations

I. Autonomous dynamical systems

By R. L. SCHAFIR
King's College, London

(*Received* 26 *February* 1981)

1. *Introduction*

When a system of differential equations is derivable from a variational principle, an outstanding consequence is the connection which ensues between symmetries and conservation laws. The idea behind the present work is to use this consequence as a new characterization of the variational principle itself. For general systems of equations there is a rather weak form of the connection, as will be demonstrated; but it has highly arbitrary and non-invariant features, and, for instance, there is not a formula associating a particular infinitesimal invariance with a particular constant of the motion. In terms of the modern technical concept which will be employed, the transformation is not *natural*. However, if the system is derivable from a variational principle, there follows a preferred transformation; and it has certain invariance properties which accord with the concept of naturality.

It will be shown that the converse is also the case: if the transformation between symmetries and conservation laws is natural (in a certain sense), then the system is derivable from a variational principle. The approach will also lead to a method for deciding whether an arbitrary system of second order differential equations is derivable from a variational principle or not, and, if it is, for calculating the different possible Lagrangians and Hamiltonians. The method is algorithmic just to the extent that it is possible to write down the general solution of the differential equations. A variant of the method will also be indicated, though more briefly, in which it is not necessary to solve the original equations, but in which instead one solves certain first order partial differential equations.

In this paper the considerations will be limited to ordinary differential equations ('dynamical systems') which are autonomous and without constraints. Later papers are planned to deal with time dependent dynamical systems and classical field theory. Here, a particular consequence of the autonomy is that the space involved is a portion of $\mathbb{R}^{2n}$, and the invariance associated with the Hamiltonian is the self-invariance of the motion field, not time translation. Also, all considerations will be entirely local, confined to regions in which there are no singularities (or zeros, when that would imply singularities) of any of the structures concerned. These qualifications should be read into the text throughout – for instance, with the use of the converse of Poincaré's Lemma.

Sections 2 and 3 will review the relevant parts of existing theory, and interpret the

0305-0041/81/0000-8880 $03.50 © 1981 Cambridge Philosophical Society

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

results in terms of a natural transformation from the constants of the motion into the infinitesimal invariances. The ideas developed will then be applied, in Section 4, to the so-called 'inverse problem' of the existence and calculation of Lagrangians and Hamiltonians for an arbitrary dynamical system. Necessary and sufficient conditions are obtained, and their use as a decision procedure demonstrated by means of some examples. Section 5 will be concerned with the naturality interpretation of these conditions, and Section 6 with the equivalence between the full theory now set up and the usual variational principle. Section 7 will conclude with a short discussion of the role of the variational principle in physical theories, in the light of the results obtained.

2. *Special and Hamiltonian vector fields*

A set of non-singular Lagrange's equations, as a set of second order ordinary differential equations, will be represented by a 'special' vector field on (a portion of) the tangent bundle $T(U)$ of a manifold, meaning a vector field of the form

$$(y^i, \eta^i(x^j, y^j)) \tag{2·1}$$

in local coordinates (x^i, y^i) $(i = 1, 2, \ldots, n)$ on $T(U) \approx V \subseteq \mathbb{R}^{2n}$.

In terms of paths, a special path is one of the form $(x^i(t), \dot{x}^i(t))$. This leads to a simple definition of speciality in index-free language: a vector field is special if and only if its integral paths are special, and a path $p: I \subseteq \mathbb{R} \to T(U)$ is special if and only if at each $(x, v_x) \in p(I)$, $(d/dt)(\pi \circ p)(t) = v_x$.

Hamilton's equations, on the other hand, are expressed in local coordinates (q^i, p_i) on what is usually interpreted as the cotangent bundle. As is well known, they may be written

$$\dot{z}^\alpha = \omega^{\alpha\beta} H_{,\beta} \quad (\alpha, \beta = 1, 2, \ldots, 2n) \tag{2·2}$$

where z^α are the coordinates (q^i, p_i), and $\omega^{\alpha\beta}$ is the matrix:

$$\omega^{\alpha\beta} = \begin{pmatrix} 0 & \mathrm{I} \\ -\mathrm{I} & 0 \end{pmatrix}, \tag{2.3}$$

and this may be made a general, tensor, equation by allowing the various terms to transform as vectors and tensors of type appropriate to their indices. It will usually be more convenient to work with the inverse $\omega_{\alpha\beta}$:

$$\omega_{\alpha\beta} = \begin{pmatrix} 0 & -\mathrm{I} \\ \mathrm{I} & 0 \end{pmatrix}, \tag{2.4}$$

so that (2.2) becomes $\omega_{\alpha\beta}\dot{z}^\beta = H_{,\alpha}$; or, in differential geometry notation:

$$M \rfloor \omega = dH, \tag{2.5}$$

where M is the vector field ('the motion').

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

The above ω is a closed 2-form of maximum rank, and so, by the inverse of Poincaré's Lemma, there exists a 1-form θ of which ω is the exterior derivative

$$\left.\begin{aligned} \omega &= d\theta \\ \text{(or in components} \quad \omega_{\alpha\beta} &= \theta_{\alpha,\beta} - \theta_{\beta,\alpha}) \end{aligned}\right\} \tag{2·6}$$

while by Darboux's theorem there exist two sets of n functions p_i and q^i on the $2n$ dimensional space, such that

$$\theta = -p_i dq^i, \tag{2.7}$$

and therefore

$$\omega = -dp_i \times dq^i. \tag{2.8}$$

Therefore, for any vector field M, if there exists a closed 2-form ω of maximum rank, and a function H, such that $M \rfloor \omega = dH$ (2·5), this equation may be regarded as a generalized form of Hamilton's equations, since there exists a coordinate system (in fact, many such), given by functions p_i, q^i of Darboux's theorem, in which ω has the skew diagonal form (2·4), and so in which (2·2) or (2·5) become the usual form of Hamilton's equations:

$$\dot{q}^i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q^i}. \tag{2·9}$$

Let us then choose for M a special vector field on the tangent bundle $T(U)$, and for ω and H a closed 2-form and a function, also on the same $T(U)$. Traditionally the Lagrangian and Hamiltonian formulations have led rather separate existences, each with its own definitions, techniques and results. But it is perfectly possible to bring the structures of the Hamiltonian formulation across to the Lagrangian formulation, by means of the Legendre transformation, allowing a Hamiltonian theory in the Lagrangian picture; much of the sequel may be regarded as doing this. Given a Lagrangian L, we have

$$q^i = x^i, \quad p_i = \frac{\partial L}{\partial y^i}, \quad H = \frac{\partial L}{\partial y^i} y^i - L, \tag{2.10}$$

and so θ and ω are given in the original 'Lagrangian coordinates' (x^i, y^i) by

$$\theta = \left(-\frac{\partial L}{\partial y^i}, \quad 0 \right), \tag{2.11}$$

$$\omega = \begin{pmatrix} \dfrac{\partial^2 L}{\partial x^i \partial y^j} - \dfrac{\partial^2 L}{\partial y^i \partial x^j} & -\dfrac{\partial^2 L}{\partial y^i \partial y^j} \\ \dfrac{\partial^2 L}{\partial y^i \partial y^i} & 0 \end{pmatrix} \tag{2.12}$$

and the equation $M \rfloor \omega = dH$ is straightforwardly equivalent to the Euler–Lagrange equations, as the reader may verify.

When we have a given motion field M, but are not starting off with a given Lagrangian for it (as will be the case), it may be best to regard the skew diagonal form

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

(2·4) as the form ω attains, not on the cotangent bundle (to which there is no intrinsic way of passing), but in certain co-ordinate systems on the same space $T(U)$, given by Darboux functions (q^i, p_i), and the Legendre transformation as a transformation from the original (x^i, y^i) to such a system.

3. *Symmetries and conservation laws*

(i) *Canonical transformations and naturality*

The established theory of the connection between symmetries and conservation laws for a Hamiltonian vector field works by using $\omega^{-1} \circ d$ as a mapping from the functions on the space to the vector fields:

$$F = \omega^{-1} \circ d\,(f). \tag{3·1}$$

Image vector fields under this mapping are called canonical. (3·1) is equivalent to

$$\pounds_F \omega = 0, \tag{3·2}$$

where £ denotes the Lie derivative. The equivalence may be shown easily from the formula

$$\pounds_F \omega = d(F \rfloor \omega) + F \rfloor d\omega, \tag{3·3}$$

for the Lie derivative of any form. (This formula will be used repeatedly in the sequel.)

Then by 'integration' of (3·2), and extension of the result to include transformations which are not members of 1-parameter groups, one may define a transformation to be canonical if and only if it preserves the 2-form

$$\omega^* = \omega. \tag{3·4}$$

The significance which will be attached here to canonical transformations, is that they form the allowable mappings under which $\omega^{-1} \circ d$ is a 'natural transformation'. The definition of a natural transformation $T: A \to B$ may be summarized as follows. If any mapping $\phi: A \to A$ (out of a certain class of allowable mappings) is applied, and a map $\phi': B \to B$, which in some sense corresponds to ϕ, is applied at the same time, then $T \circ \phi = \phi' \circ T$. Thus T is invariant under the class of allowable mappings; the idea is that it is independent of any extraneous structure on A and B which would be altered by the application of ϕ, ϕ'.

In the present case the two structures are the set of functions and the set of vector fields, both on $V \subseteq \mathbb{R}^{2n}$. The sense in which the mappings correspond is that they are both the induced map of an underlying map $\phi: V \to V$. Thus we have that $\omega^{-1} \circ d$ commutes with ϕ_* and ϕ^{-1*}, for any allowable (non-singular) mapping $\phi: V \to V$. This is summarized by the commutation diagram (Fig. 1).

Given f, and $F = \omega^{-1} \circ df$, we have

$$\phi^{-1*}(F \rfloor \omega) = \phi^{-1*}(df) = d(\phi^{-1*} f), \quad \text{from} \quad d\phi^* = \phi^* d.$$

$$= (\phi_* F) \rfloor \omega, \quad \text{from the diagram (Fig. 1).}$$

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

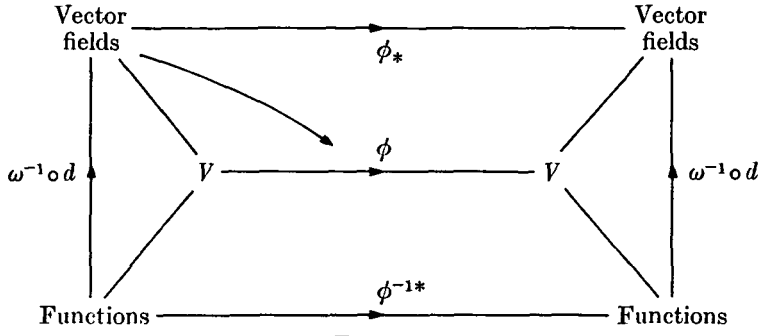


Fig. 1

$$\left(\text{In components } \frac{\partial x^\beta}{\partial \phi^\alpha}\,\omega_{\beta\gamma}\,F^\gamma = \frac{\partial \phi^\beta}{\partial x^\gamma}\,F^\gamma\,\omega_{\alpha\beta}\right)$$

Therefore

$$(\phi_* F) \rfloor (\phi^{-1*}\omega) = (\phi_* F) \rfloor \omega. \tag{3·5}$$

for all $F \in im\,(\omega^{-1} \circ d)$, and so, by the non-singularity of ϕ

$$\phi^{-1*}\omega = \omega \quad \text{or} \quad \phi^*\omega = \omega,$$

which is the formula (3·4). Hence the class of allowable ϕ are the canonical transformations.

But these are precisely the transformations generated by images of functions under $\omega^{-1} \circ d$. Figure 1 commutes under its 'own' transformations, these being the transformations generated by $\omega^{-1} \circ d$ itself. This is indicated in the diagram by the arrow from the top left-hand corner.

For a physical illustration of the idea, consider two inertial observers, in some theory, with a transformation from constants of the motion to generators of invariances, given by an ω with this property. Observer A associates (say) translation in his time direction with his expression for the energy, his x-translation with his x-momentum, and so on. Now transform to inertial observer B, i.e. transform by one of the invariances of the theory. Then B associates *his* time translation with *his* expression for the energy, *his* x-translation with *his* x-momentum, and so on. The transformation between symmetries and conservation laws is itself invariant under those self-same invariances.

This 'self-invariance' naturality property is characteristic of a closed 2-form. If ω is a closed 2-form, we have the equivalence of the two definitions of canonical transformation:

$$\exists f \; s.t. \; F \rfloor \omega = df \quad \text{if and only if} \quad \pounds_F \omega = 0,$$

and the property is simply a statement of this. Conversely, if ω^{-1} is a non-singular mapping from 1-forms to vector fields, linear at each point, and has the above naturality property, then ω is a closed 2-form. This converse will be demonstrated in Section 6.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

The diagram (Fig. 1) states that

$$\omega^{-1} \circ d(\phi^{-1*} f) = \phi_* F, \tag{3·7}$$

of which the infinitesimal version is

$$\omega^{-1} \circ d(\pounds_G f) = \pounds_G F, \tag{3.8}$$

i.e. $\omega^{-1} \circ d$ preserves Lie derivatives (with respect to any member of its own image). In these terms one can summarize speedily the way in which constants of the motion correspond to canonical infinitesimal invariances. If f is a constant of the motion M, F the corresponding canonical vector field,

$$\pounds_F M = -\pounds_M F = -\pounds_M(\omega^{-1} \circ df) = -\omega^{-1} \circ d(\pounds_M f) = 0, \tag{3.9}$$

and so F is the infinitesimal generator of a 1-parameter group of invariances ('infinitesimal invariance' for short). The converse is almost – but not quite – true: if $\pounds_M F = 0$, then $\pounds_M f =$ constant; the rate of change of f along M is a constant, but not necessarily zero. This slight complication will have to be taken into account; let us call functions such that $\pounds_M f =$ constant 'up-to-a-factor constants of the motion', and use the term 'constants' to include both these and the 'proper' constants for which $\pounds_M f = 0$.

(ii) *Point transformations*

In the Lagrangian picture it is usual to confine the class of transformations under consideration to the following type:

$$(x, v_x) \to (\phi(x), \phi_*(v_x)), \quad \text{for} \quad (x, v_x) \in T(U) \tag{3.10}$$

known as point transformations.

In Lagrangian co-ordinates (x^i, y^i), (3·10) is

$$(x^i, y^i) \to \left(\bar{x}^i(x^j), \frac{\partial \bar{x}^i}{\partial x^j} y^j\right), \tag{3·11}$$

and the infinitesimal version is vector fields of the form

$$\left(\xi^i(x^j), \frac{\partial \xi^i}{\partial x^j} y^j\right). \tag{3.12}$$

Point transformations transform any special path or vector field into another special path or vector field.

It was between certain infinitesimal point transformations and constants of the motion that the classical result – Noether's theorem – was obtained. Let us look at Noether's theorem within the present context.

We have a special vector field derived from a Lagrangian, and hence a θ and an ω, given by (2·11) and (2·12). For an infinitesimal point transformation X, $= (\xi^i, (\partial \xi^i / \partial y^j) y^j)$, to be associated with some function, it must, by definition, be canonical, and so

$$\pounds_X \omega = 0 \quad \text{or} \quad \pounds_X \theta = d\phi, \quad \text{for some function } \phi. \tag{3.13}$$

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

Calculating in components:

$$\pounds_X\left(-\frac{\partial L}{\partial y^i},\ 0\right) = -\left(\frac{\partial^2 L}{y^i\,\partial x^j}\xi^j + \frac{\partial^2 L}{\partial y^i\,\partial y^j}\frac{\partial \xi^j}{\partial x^k}y^k + \frac{\partial L}{\partial y^j}\frac{\partial \xi^j}{\partial x^i},\ 0\right)$$

$$= \left(-\frac{\partial \pounds_X L}{\partial y^i},\ 0\right) = \left(\frac{\partial \phi}{\partial x^i},\ \frac{\partial \phi}{\partial y^i}\right),$$

Therefore

$$\pounds_X L = \frac{\partial \phi(x^j)}{\partial x^i}y^i + \gamma(x^j). \tag{3.14}$$

Thus a point transformation is canonical if and only if it preserves L up to a total time derivative plus a function of the x^i. Further, $\gamma(x^i)$ = constant for X an invariance, because

$$-\pounds_M f = \pounds_X H = \pounds_X\left(\frac{\partial L}{\partial y^i}y^i - L\right) = \frac{\partial \pounds_X L}{\partial y^i}y^i - \pounds_X L \tag{3.15}$$

and this is a constant if and only if $\gamma(x^i)$ = constant.

The function f associated with X is given by

$$df = X \rfloor \omega = \pounds_X \theta - d(X \rfloor \theta), \quad \text{from (3·3)}$$

$$= \operatorname{grad}\left(-\phi + \frac{\partial L}{\partial y^i}\xi^i\right), \tag{3·16}$$

which gives us the formula of Noether's theorem (for transformations not involving the time).

We see why Noether's theorem does not deal with all point invariances, but is confined to the subclass which preserve the Lagrangian up to a total time derivative plus a constant – it is just these which are canonical. Although it is customary to speak of the transformation as 'between symmetries and conservation laws', actually it goes strictly one way: from constants of the motion into, but not onto, the infinitesimal invariances. We see also that there is only a very partial converse to Noether's theorem; any constant of the motion gives rise to an infinitesimal invariance, but it need not be a point transformation, and, if it is not, it need not preserve the Lagrangian in any way.

(iii) *The non-natural transformation*

Suppose we have only a special vector field M† on V, without any other structure. Its solutions

$$\bar{x}^\alpha = f^\alpha(t, x^\beta), \tag{3.17}$$

are a flow of curves in the space (one curve through each point), and form a 1-parameter Lie group of transformations of the space into itself, $x^\alpha \to \bar{x}^\alpha$, with t the canonical parameter.

† For this section M does not need to be special, but the application will only be to special vector fields.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

By a fundamental theorem, there exists for any vector field such as M a set of 'comoving' coordinates h^α in which M has the form $\partial/\partial h^1$. The h^α are a maximal set of functionally independent functions on V, and their gradients span a $2n$ dimensional vector space of 1-forms linearly independent at each point.

The h^α may be pictured as made up of $2n-1$ arbitrary functionally independent functions defined on an initial hypersurface crossing the curves, each extended to a function on the whole space by defining it to be constant along each curve; plus one further function, arbitrary on the initial hypersurface, and with rate of change a constant (here, unity) along each curve.

Similarly, given $M = \partial/\partial h^1$, we may choose any $2n-1$ vector fields, linearly independent at each point, whose components do not depend on h^1, and add M itself to make up a $2n$ dimensional vector space of infinitesimal invariances. Geometrically, this is to choose any $2n-1$ vector fields, linearly independent at each point, lying within the initial hypersurface, and drag them along the motion field M; then add the self-invariance formed by M.

Thus for a general system we can construct the vector space of gradients of up-to-a-factor constants of the motion, and a vector space of infinitesimal invariances. These two are real vector spaces of the same dimension, and hence isomorphic. But the isomorphism is not natural, for to set it up depends on an arbitrary choice of basis in each space, and it will not be invariant under any particular class of mappings.

4. *Derivability from a Lagrangian*

Given a special vector field $M = (y^i, \eta^i(x^j, y^j))$, we approach the problem of constructing a Lagrangian formulation for it via a Hamiltonian formulation on the same space, i.e. we seek to introduce a closed 2-form ω and a function H, all on the same tangent bundle $T(U) \approx V \subseteq \mathbb{R}^{2n}$, such that $M \rfloor \omega = dH$ (2·5).

Choose co-moving coordinates h^α so that $M = \partial/\partial h^1$. In these coordinates (2.5) becomes

$$\omega_{\alpha 1} = H_{,\alpha}. \tag{4·1}$$

Then since

$$\omega_{\alpha 1} = \theta_{\alpha,1} - \theta_{1,\alpha},$$

we need only choose a θ all of whose components are independent of h^1; i.e. choose any $2n$ components $\theta_\alpha(h^2, h^3, \ldots, h^{2n})$. Then $\omega_{\alpha 1}$ is automatically a gradient, and (4·1) and hence (2·5) are satisfied. In coordinate free language, we have shown that there always exists a θ such that $\pounds_M \theta = 0$; and

$$\left.\begin{aligned} 0 = \pounds_M \theta = d(M \rfloor \theta) + M \rfloor d\theta \\ H = -(M \rfloor \theta). \end{aligned}\right\} \tag{4·2}$$

so that

Thus any vector field, including any special vector field can be made Hamiltonian, in the generalized sense of that term that there exists an ω and an H such that $M \rfloor \omega = dH$. Does that mean that any special vector field, representing a system of second order ordinary differential equations, may be derived from the Euler–Lagrange

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

equations of some Lagrangian? The answer is no, and to see why, let us examine the procedure by which one would hope to pass from the above Hamiltonian formulation to a Lagrangian formulation, and find where a difficulty arises.

Given M, ω and H, there exist p_i and q^i satisfying Darboux's theorem (2·7, 2·8); suppose that we have found these functions. Transform M and ω by the transformation $(x^i, y^i) \to (q^i, p_i)$, so that $M \rfloor \omega = dH$ becomes the usual Hamilton equations (2·9). Now apply the usual inverse Legendre transformation: define a further set of functions $\dot{q}^i$, and transform again by $(q^i, p_i) \to (q^i, \dot{q}^i)$. $\dot{q}^i$ may be defined either by $\dot{q}^i = \partial H/\partial p_i$, or $\dot{q}^i = \pounds_M q^i$; the latter may be characterized as that transformation, holding the tangent spaces fixed, which transforms the vector field once more into a special vector field. We may now define a Lagrangian in the usual way:

$$L = p_i \dot{q}^i - H. \tag{4·3}$$

But we have transformed the vector field M by $(x^i, y^i) \to (q^i, \dot{q}^i)$, and this in general changes M. (The theorem that any vector field can be transformed to $\partial/\partial h^1$ shows that vector fields which differ from each other by a transformation must be regarded as essentially different.) We have derived a Lagrangian for what is in general a different system of differential equations to the one we started out with.

This argument may, if desired, be rephrased in terms involving the cotangent bundle. Choose an ω and an H on the tangent bundle $T(U)$ so that $M \rfloor \omega = dH$ (this is possible by the integration theorem for ordinary differential equations). Next, choose a transformation to the cotangent bundle $T^*(U)$ which pulls the canonical 2-form back to ω. (Such transformations exist by Darboux's theorem.) This takes M to a Hamiltonian vector field M_* on $T^*(U)$. Finally, apply the inverse Legendre transformation to take M_* to a special vector field M_{**} back on $T(U)$. We have a Lagrangian defined for M_{**}; but in general $M_{**} \neq M$.

For example, consider the 1-dimensional harmonic oscillator, expressed by the special vector field $(y, -x)$. A choice of comoving coordinates h_1, h_2† is

$$h_1 = \tan^{-1}\left(\frac{x}{y}\right), \quad h_2 = \tfrac{1}{2}(x^2 + y^2).$$

Choose $\theta = (h_2, 0)$, say.

Now with Darboux functions associated with a θ for which $\pounds_M \theta = 0$, the transformation $(x^i, y^i) \to (q^i, \dot{q}^i)$ is singular.

(For if $$0 = \pounds_M(p_i\, dq^i) = (\pounds_M p_i)\, dq^i + p_i\, d\pounds_M q^i = (\pounds_M p_i)\, dq^i + p_i\, d\dot{q}^i, \tag{4.4}$$

so we have a linear relationship between the gradients of the functions q^i, $\dot{q}^i$.)
So let us add to the above a term $d\phi$, with $\pounds_M \phi \neq 0$. Here let us take, for example, $\phi = h_1 h_2$. The resulting θ is $\theta = (2h_2, h_1)$. This $\theta = -(p\, \partial q/\partial h_1, p \partial q/\partial h_2)$, where $p = -h_1 h_2$, $q = \log_e(h_1^2\, h_2)$, as the reader may verify.

Transform the vector field to (q, p) coordinates, where it becomes

$$M = (-2p\, e^{-q}, -p^2 e^{-q}).$$

† In examples, coordinate indices will be written as subscripts.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

Introduce $\dot{q} = \pounds_M q = -2p\,e^{-q}$. Transform the vector field again, to the $(q, \dot{q})$ system. It becomes $(\dot{q}, -\frac{1}{2}\dot{q}^2)$.

$$H = -h_2 = -p^2 e^{-q} = -\tfrac{1}{4}\dot{q}^2 e^q, \quad \text{and so} \quad L = p\frac{\partial H}{\partial p} - H = -\tfrac{1}{4}\dot{q}^2 e^q.$$

Thus we have obtained a Lagrangian, but in terms of q and $\dot{q}$, for the vector field $(\dot{q}, -\frac{1}{2}\dot{q}^2)$, which is not of the same functional form as $(y, -x)$. Equally, if one pulls back the Lagrangian to the original coordinate system, one obtains $L = -\frac{1}{2}(x^2+y^2)$, which is not a Lagrangian for $(y, -x)$.

Let us look for conditions that the special vector field remain invariant under $(x^i, y^i) \to (q^i, \dot{q}^i)$. An obvious sufficient condition is that this transformation be the identity; since $\dot{q}^i = \pounds_M q^i$, it is enough for this that

$$q^i = x^i. \tag{4·5}$$

And this condition is also necessary. For, if M is invariant under $(x^i, y^i) \to (q^i, \dot{q}^i)$, and so is derivable from a Lagrangian, then the standard theory gives $q^i = x^i$, $p_i = \partial L/\partial y^i$ for the Darboux functions, satisfying (4.5).

So given a 1-form θ constructed to make M locally Hamiltonian, we want another 1-form $\bar{\theta}$, leading to the same $\omega = d\theta$ and of the form $-\bar{p}_i\,dx^i$. θ and $\bar{\theta}$ must therefore differ by a gradient

$$\bar{\theta} = \theta + d\phi.$$

Thus

$$-\bar{p}_i dx^i = -p_i\,dq^i + d\phi, \tag{4·6}$$

i.e. we seek a function ϕ which generates a canonical transformation $(q^i, p_i) \to (x^i, \bar{p}_i)$. In terms of this ϕ there is a convenient expression for the Lagrangian, namely

$$L = -\pounds_M \phi. \tag{4·7}$$

To obtain this result, calculate $\pounds_M(-p_i,\ 0) = \pounds_M(\theta + d\phi) = d\pounds_M \phi$ in (x^i, y^i) coordinates; one obtains

$$\left(\frac{\partial p_i}{\partial x^j}\,y^j + \frac{\partial p_i}{\partial y^j}\,\eta^j,\ p_i\right) = -\left(\frac{\partial \pounds_M \phi}{\partial x^i},\ \frac{\partial \pounds_M \phi}{\partial y^i}\right). \tag{4·8}$$

The second set of n equations are defining equations for p_i, and then the first n are the Euler–Lagrange equations.

To be able to add a $d\phi$ which makes the y^i components of $\bar{\theta}$ zero, the y^i components of θ, as functions of y^i (holding the x^i fixed) must satisfy the 'curl-free conditions'

$$\frac{\partial \theta_{y^i}}{\partial y^j} - \frac{\partial \theta_{y^j}}{\partial y^i} = 0, \tag{4·9}$$

i.e.

$$\omega_{y^i y^j} = 0. \tag{4·10}$$

In one degree of freedom the curl-free conditions are satisfied identically, but in more than one, they amount to a real restriction.

To sum up: the special vector field M can be derived from a Lagrangian if and only if

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

there exists a 2-index object ω which satisfies the algebraic conditions

$$\left.\begin{aligned} &\text{(i)} \quad \omega_{\alpha\beta} = -\omega_{\beta\alpha}. \\ &\text{(ii)} \quad \det\omega \neq 0 \text{ (non-singularity condition).} \\ &\text{(iii)} \quad \omega_{y^i y^j} = 0 \text{ (curl-free conditions).} \end{aligned}\right\} \tag{4.11}$$

and the differential conditions

$$\left.\begin{aligned} &\text{(iv)} \quad \pounds_M \omega = 0. \\ &\text{(v)} \quad d\omega = 0. \end{aligned}\right\}$$

A number of previous authors have given equations which can be interpreted as equations for components of $\omega_{\alpha\beta}$, equivalent to equations (4·11) above. (For a compendium of previous work on this question, see Santilli (2).)

In comoving coordinates h^α, one may express the conditions in terms of the 1-form θ:

$$\left.\begin{aligned} &\text{(i)} \quad \theta_\alpha = \theta_\alpha(h^a), \quad \text{for} \quad a \neq 1. \\ &\text{(ii)} \quad \det\omega \neq 0. \\ &\text{(iii)} \quad \frac{\partial h^\alpha}{\partial y^i}\frac{\partial h^\beta}{\partial y^j}\omega_{\alpha\beta} = 0, \end{aligned}\right\} \tag{4.12}$$

where $\omega = d\theta$.

Equation (4·12) may be used as a decision procedure, algorithmic to within certain simple integrals, as shown by the following examples.

(A) 1-*dimensional harmonic oscillator* $(y, -x)$

A pair of comoving coordinates (already introduced) is

$$h_1 = \tan^{-1}\left(\frac{x}{y}\right), \quad h_2 = \tfrac{1}{2}(x^2+y^2).$$

Because the curl-free conditions are always satisfied, this, and any system in 1 degree of freedom, is always derivable from a Lagrangian.

The 'orthodox' choice for θ is $\theta_1 = -h_2$, $\theta_2 = 0$. Hence $H = h_2$. Transforming to (x, y) coordinates, $\theta = (\frac{1}{2}y, -\frac{1}{2}x)$, and we want a function ϕ so that $(\frac{1}{2}y, -\frac{1}{2}x) + \phi_{,\alpha}$ has y component equal to zero. Therefore $\phi = \int \frac{1}{2}x dy = \frac{1}{2}xy$, and therefore, from

$$L = py - H \quad \text{or} \quad L = -\pounds_M \phi,$$

we have $L = \frac{1}{2}(x^2 - y^2)$.

Suppose we chose, instead, $\theta_1 = h_2^2$, $\theta_2 = h_2$ (so that $H = -h_2^2$). Then we get $\theta = \frac{1}{2}(x^2+y^2)(x+\frac{1}{2}y, y-\frac{1}{2}x)$, in (x, y) coordinates. We need a ϕ so that

$$\tfrac{1}{2}(x^2+y^2)(y-\tfrac{1}{2}x) + (\partial\phi/\partial y) = 0.$$

Integrate for ϕ, and from $L = -\pounds_M \phi$, $L = \frac{1}{12}y^4 + \frac{1}{2}x^2y^2 - \frac{1}{4}x^4 + \frac{1}{2}x^3y$; and the reader may check that this is indeed a Lagrangian for $(y, -x)$. We could obviously get as many different Lagrangians as we like, by making different choices of θ; conversely, every Lagrangian can be made to arise in this way.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

(B) *Falling body over flat earth, in 2 dimensions*

The special vector field is $(y_1, y_2, 0, -1)$. (The gravitational $g = 1$, for convenience.) There are the following expressions for a set of comoving coordinates in terms of the Lagrangian coordinates, and vice versa:

$$\begin{aligned} h_1 &= -y_2 & x_1 &= h_4 + h_1 h_2 \\ h_2 &= y_1 & x_2 &= h_3 - \tfrac{1}{2}h_1^2 \\ h_3 &= \tfrac{1}{2}y_2^2 + x_2 & y_1 &= h_2 \\ h_4 &= x_1 + y_1 y_2 & y_2 &= -h_1. \end{aligned}$$

Suppose we have chosen a 1-form $\theta_\alpha(h_2, h_3, h_4)$. There is just one curl-free condition: $(\partial h_\alpha/\partial y_1)(\partial h_\beta/\partial y_2)\omega_{\alpha\beta} = 0$, and working this out, we get

$$h_1^2(\omega_{24}) + h_1(\omega_{14} + \omega_{23}) - (h_2\omega_{24} + \omega_{12}) = 0$$

Each term in brackets does not involve h_1, but the curl-free conditions hold for all h_1, and so the brackets must vanish separately:

$$\omega_{23} = -\omega_{14} = -\frac{\partial\theta_1}{\partial h_4}, \quad \omega_{34} = 0, \quad \omega_{24} = -\frac{1}{h_2}\frac{\partial\theta_1}{\partial h_2}.$$

This allows us to fill in some of the entries in $\omega_{\alpha\beta}$:

$$\omega_{\alpha\beta} = \begin{pmatrix} 0 & \dfrac{\partial\theta_1}{\partial h_2} & \dfrac{\partial\theta_1}{\partial h_3} & \dfrac{\partial\theta_1}{\partial h_4} \\ -\dfrac{\partial\theta_1}{\partial h_2} & 0 & \dfrac{\partial\theta_1}{\partial h_4} & -\dfrac{1}{h_2}\dfrac{\partial\theta_1}{\partial h_2} \\ -\dfrac{\partial\theta_1}{\partial h_3} & -\dfrac{\partial\theta_1}{\partial h_4} & 0 & 0 \\ -\dfrac{\partial\theta_1}{\partial h_4} & \dfrac{1}{h_2}\dfrac{\partial\theta_1}{\partial h_2} & 0 & 0 \end{pmatrix}.$$

$H = -\theta_1$. The orthodox choice is now $\theta_1 = -\frac{1}{2}h_2^2 - h_3$. But let us look for other possible choices. Could we take the horizontal momentum h_2 for the Hamiltonian? We can see from the above matrix that the choice $\theta_1 = -h_2$ leads to the third row becoming all zeros, which violates non-singularity. Therefore this is not a possible choice.

However $\theta_1 = h_4$ is a possible choice; it leads to

$$\omega_{\alpha\beta} = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

In this case the 1-form can be obtained more or less by inspection; for more general cases there is the formula

$$\theta = \int_0^1 \phi_t^*(\dot{\phi}_t \lrcorner\, \omega)\, dt, \tag{4.13}$$

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

where ϕ_t is the map

$$\phi_t(x^i) = tx^i. \tag{4.14}$$

(See Lang (1), p. 134.)

So, for the present example,

$$\theta = (-h_4,\ h_3h_4,\ h_3h_4 - h_2,\ h_2h_3)$$

Transforming it to (x_i, y_i) coordinates:

$$\theta = (\tfrac{1}{2}y_1y_2^2 + y_1x_2,\ x_1y_1 + y_1^2y_2 - y_1,$$
$$x_1x_2 + \tfrac{1}{2}x_1y_2^2 + y_1y_2^3 + 2y_1y_2x_2,\ x_1 + y_1y_2x_1 + \tfrac{3}{2}y_1^2y_2^2 + y_1^2x_2).$$

By construction there exists a ϕ such that the last two components are $\partial\phi/\partial y_1$, $\partial\phi/\partial y_2$; after more calculation

$$\phi = x_1y_2 + x_1x_2y_1 + \tfrac{1}{2}x_1y_1y_2^2 + \tfrac{1}{2}y_1^2y_2^3 + y_1^2y_2x_2.$$

Finally, from $L = -\pounds_M\phi$, we have $L = x_1 - y_1y_2$.

In this variational principle for $(y_1, y_2, 0, -1)$, *vertical* translation is associated with *horizontal* momentum, while horizontal translation is associated with the up-to-a-factor (but not proper) constant of the motion $-h_1 = y_2$.

Is it merely an accident of the particular choice of comoving coordinates that the horizontal momentum h_2 cannot be chosen for Hamiltonian? It is straightforward to demonstrate the following:

(i) Any transformation to a different system of comoving coordinates $\bar{h}^\alpha$ must be of the form

$$\bar{h}^1 = h^1 + \bar{h}(h^a), \quad \bar{h}^a = \bar{h}^a(h^b), \quad \text{for} \quad a, b \neq 1.$$

(ii) Any such transformation leaves θ_1 unchanged.

(iii) Therefore, if a choice of θ_1 makes ω singular in h-coordinates, it remains singular in $\bar{h}$-coordinates. Hence it is an invariant fact that the horizontal momentum cannot be the Hamiltonian of any Lagrangian for this system.

The method so far requires knowledge of a system of comoving coordinates, which is equivalent to a knowledge of the general solution of the equations. For an alternative approach, as mentioned in the introduction, work in the same (x^i, y^i) coordinates throughout, and solve, for the components of ω, the partial differential equations expressed by (4.11), viz.

(i) $\pounds_M\omega = 0$ (equivalent to $M \lrcorner \omega = dH$),

(ii) $d\omega = 0$,

together with the conditions

$$\omega_{\alpha\beta} = -\omega_{\beta\alpha}, \quad \det\omega \neq 0 \quad \text{and} \quad \omega_{y^iy^j} = 0.$$

For this example let us try $\omega_{12} = 0$, $\omega_{13} = x_1 + y_1y_2$. The other $d\omega = 0$ equations are satisfied by

$$\omega_{14} = \tfrac{1}{2}y_1^2 + \gamma(x_2, y_2), \quad \text{with} \quad \frac{\partial\gamma}{\partial x_2} = \frac{\partial\omega_{24}}{\partial x_1}, \quad \frac{\partial\gamma}{\partial y_2} = \frac{\partial\omega_{24}}{\partial y_1}.$$

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

Let $\gamma = 0$, $\omega_{24} = 1$. Then all equations are satisfied, and

$$\omega_{\alpha\beta} = \begin{pmatrix} 0 & 0 & x_1 + y_1 y_2 & \frac{1}{2}y_1^2 \\ 0 & 0 & \frac{1}{2}y_1^2 & 1 \\ -x_1 - y_1 y_2 & -\frac{1}{2}y_1^2 & 0 & 0 \\ -\frac{1}{2}y_1^2 & -1 & 0 & 0 \end{pmatrix}$$

The formulae (4·13), (4·14) give for the 1-form:

$$\theta = (\tfrac{1}{3}x_1 y_1 + \tfrac{3}{8}y_1^2 y_2,\ \tfrac{1}{8}y_1^3 + \tfrac{1}{2}y_2,\ -\tfrac{1}{3}x_1^2 - \tfrac{1}{4}x_1 y_1 y_2 - \tfrac{1}{8}x_2 y_1^2,\ -\tfrac{1}{8}x_1 y_2^2 - \tfrac{1}{2}x_2)$$

The last two components are $\partial\phi/\partial y_1$, $\partial\phi/\partial y_2$, and so

$$\phi = -\tfrac{1}{3}x_1^2 y_1 - \tfrac{1}{8}x_1 y_1^2 y_2 - \tfrac{1}{24}x_2 y_1^3 - \tfrac{1}{2}x_2 y_2.$$

The θ constructed in this approach does not satisfy $\pounds_M \theta = 0$, so we cannot use the formula $L = -\pounds_M \phi$. Instead we proceed by finding p_i and H.

$$\overline{\theta} = (-p_1,\ -p_2,\ 0, 0) = (x_1 y_1 + \tfrac{1}{2}y_1^2 y_2,\ \tfrac{1}{6}y_1^3 + y_2, 0, 0)$$

$$dH = M \lrcorner \omega = (-\tfrac{1}{2}y_1^2,\ -1,\ -x_1 y_1 - \tfrac{3}{2}y_1^2 y_2,\ -\tfrac{1}{2}y_1^3 - y_2).$$

Therefore

$$H = -\tfrac{1}{2}x_1 y_1^2 - x_2 - \tfrac{1}{2}y_1^3 y_2 - \tfrac{1}{2}y_2^2,$$

and so

$$L = p_i y^i - H = x_2 - \tfrac{1}{2}x_1 y_1^2 - \tfrac{1}{6}y_1^3 y_2 - \tfrac{1}{2}y_2^2.$$

For this Lagrangian horizontal translation is not even canonical (and so is not associated with any constant), as can be seen by direct use of $\omega_{\alpha\beta}$, or by noting that L does not remain invariant to within a total time derivative plus a constant (c.f. Section 3, part ii).

(C) *A non-Lagrangian example*

The h-function approach in example B also shows how to construct a system which is provably *not* derivable from a Lagrangian. In B the curl-free condition was used to provide three algebraic equations connecting the components of $\omega_{\alpha\beta}$. We look for an example where the number of such equations is greater than, or equal to, the number of independent components of ω.

We had three equations essentially because the h_α involved at most quadratic terms in the x_i and y_i. $\omega_{\alpha\beta}$ has six algebraically independent components, so let us choose, e.g.:

$$x_1 = h_4 + \tfrac{1}{2}h_1^2 h_2, \quad x_2 = h_3 + \tfrac{1}{5}h_1^5.$$

Speciality requires that $y_i = \partial x_i/\partial h_1$; so the full system has to be:

$$\begin{aligned} x_1 &= h_4 + \tfrac{1}{2}h_1^2 h_2 & h_1 &= y_2^{\frac{1}{4}} \\ x_2 &= h_3 + \tfrac{1}{5}h_1^5 & h_2 &= y_1 y_2^{-\frac{1}{4}} \\ y_1 &= h_1 h_2 & h_3 &= x_2 - \tfrac{1}{5}y_2^{\frac{5}{4}} \\ y_2 &= h_1^4 & h_4 &= x_1 - \tfrac{1}{2}y_1 y_2^{\frac{1}{4}}. \end{aligned}$$

Hence the special vector field is $(y_1, y_2, y_1 y_2^{-\frac{1}{4}}, 4y_2^{\frac{3}{4}})$, in the (x_i, y_i) coordinates.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

The curl-free condition $(\partial h_\alpha/\partial y_1)(\partial h_\beta/\partial y_2)\omega_{\alpha\beta} = 0$ gives

$$h_1^{-4}(\tfrac{1}{4}\omega_{21}) + h_1^{-3}(-\tfrac{1}{8}\omega_{24} + \tfrac{1}{8}\omega_{42}) + h_1^{-2}(-\tfrac{1}{8}\omega_{41}) + (-\tfrac{1}{4}\omega_{24}) + h_1^{2}(\tfrac{1}{8}\omega_{43}) = 0,$$

and so $\omega_{21} = \omega_{41} = \omega_{23} = \omega_{43} = 0$. $\omega_{24} = +\omega_{42}$ and so $\omega_{24} = 0$, i.e. $\omega_{41} = \omega_{42} = \omega_{43} = \omega_{44} = 0$; $\omega_{\alpha\beta}$ is singular for any choice of θ_1 in this system of h-coordinates, and hence, as before, in any system of h-coordinates.

This shows that the special vector field, i.e. the pair of differential equations

$$\ddot{x}_1 - \dot{x}_1 \dot{x}_2^{-\frac{1}{4}} = 0,$$

$$\ddot{x}_2 - 4\dot{x}_2^{\frac{3}{4}} = 0,$$

cannot be derived from any Lagrangian.

5. *Interpretation of the curl-free conditions*

Let us now look for an invariance-type interpretation of the curl-free conditions $\omega_{y^i y^j} = 0$ (4·10). It will be convenient to work with the equivalent expression for ω^{-1}:

$$\omega^{x^i x^j} = 0, \tag{5·1}$$

(5·1) is preserved by any point transformation, as is easy to check. Also, point transformations are the transformations which preserve the Lagrangian formulation (analogous to the role of canonical transformations in the Hamiltonian picture), that is, they have the property that the transformed Lagrangian is a Lagrangian for the transformed equations. This suggests that we want some aspect of the transformation from constants of the motion to infinitesimal invariances which is invariant under point transformations. When that transformation is limited to infinitesimal point invariances and their associated constants, then there is the formula of Noether's theorem

$$f = -\phi + \frac{\partial L}{\partial y^i}\xi^i.$$

We will look for something associated with this formula, but which does not involve L explicitly.

Now in the standard physical theories, the constants of the motion are not just a set of scalars, but are grouped together to make up higher order geometric objects. This is a feature which has been ignored in the account so far, but which arguably should play a part in any complete theory of conservation laws. For instance, the components of momentum group together to form a vector (and in field theory energy and momentum form a tensor), and the components of angular momentum form a skew-symmetric 2-index tensor. Let us use the terminology 'conserved quantity' for such a geometric object, with the term 'constant of the motion' meaning, as before, an individual component. Each conserved quantity corresponds to a set of infinitesimal canonical point invariances which group together in the same way, in configuration space, and form

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

geometric objects of one higher order. If the conserved quantity f_A transforms under certain point transformations by

$$\bar{f}_A = S_A{}^B f_B, \quad \text{with} \quad S_A{}^B = S_A{}^B(x^j) \tag{5·2}$$

the configuration space part of the corresponding invariance object transforms by

$$\bar{\xi}^i{}_A = \frac{\partial \bar{x}^i}{\partial x^j} S_A{}^B \xi^j{}_B \tag{5·3}$$

We will see that the curl-free conditions are equivalent to the preservation of this property under a certain class of (non-canonical) point transformations, i.e.

$$\left.\begin{aligned} &\text{if} && f_A \xrightarrow{\omega^{-1} \circ d} \left(\xi^i{}_A, \quad \frac{\partial \xi^i{}_A}{\partial x^j} y^j\right) \\ &\text{then} && \bar{f}_A \xrightarrow{\omega^{-1} \circ d} \left(\bar{\xi}^i{}_A, \quad \frac{\partial \bar{\xi}^i{}_A}{\partial \bar{x}^j} \bar{y}^j\right). \end{aligned}\right\} \tag{5·4}$$

This is not a naturality condition as it stands, since ω^{-1} is also transformed by the mapping, but the imposition of further structure for $\omega^{-1} \circ d$ to preserve in the role of an isomorphism. Naturality is involved, rather, in the property preserved, that the conserved objects group together to make up higher order objects. The higher order object is more 'natural' than the components, for the individual components depend on various bits of extraneous data, such as a choice of direction, but the higher order object loses such dependence.

One can explain how the infinitesimal invariances can be put together to make up higher order objects, in terms of the group structure.† Suppose we have an r-parameter Lie group of transformations of n-space into itself, expressed in coordinates by

$$\left.\begin{aligned} x^i \to \bar{x}^i = f^i(u^A, x^j); \quad (i, j = 1, 2, \ldots, n), \\ (A = 1, 2, \ldots, r). \end{aligned}\right\} \tag{5·5}$$

By group closure there exists a function $\phi^A(u^B, u^C)$ such that

$$f^i(u_2{}^A, f^j(u_1{}^B, x^k)) = f^i(\phi^A(u_1{}^B, u_2{}^C), x^j) \tag{5·6}$$

By definition the infinitesimal generators $\xi^i{}_A$ are

$$\xi^i{}_A = \frac{\partial f^i}{\partial u^A}(id., x^j) \tag{5.7}$$

and the left and right transformation functions are

$$\chi^A{}_B = \frac{\partial \phi^A}{\partial u_1{}^B}(id., u_2{}^C), \quad \psi^A{}_B = \frac{\partial \phi^A}{\partial u_2{}^B}(u_1{}^C, id.) \tag{5·8}$$

By differentiating (5·6) with respect to each variable, and setting each equal to the

† The argument of the next few paragraphs presumably exists somewhere in the literature, but I have not found it.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

identity, we obtain the following result, which may be regarded as a form of Lie's first theorem:

$$\xi^i{}_A(\bar{x}^j) = \frac{\partial \bar{x}^i}{\partial x^j} S_A{}^B \xi^j{}_B, \tag{5·9}$$

where $S_A{}^B = \chi^{-1C}{}_A \psi^B{}_C$, a function of the u^A alone.

Thus *if* we define $\xi^i{}_A$ to be the components of an object whose transformation law is

$$\bar{\xi}^i{}_A = \frac{\partial \bar{x}^i}{\partial x^j} S_A{}^B \xi^j{}_B \tag{5.10}$$

then $\xi^i{}_A$ is invariant under the group. If we can extend the class of transformations for which (5·9) can be written, with $S_A{}^B(x^j)$ reducing to $\chi^{-1C}{}_A \psi^B{}_C$, for members of the original group, then $\xi^i{}_A$ becomes a more general object which is invariant under the group which it itself generates.

The transformation law is extendible if, for example, one starts with a linear or affine group (which covers the elementary groups of point invariances which occur in physical theories). Suppose

$$\bar{x}^i = A^i{}_j(u^A)\, x^j. \tag{5.11}$$

Let $a^i{}_{Aj} = \dfrac{\partial A^i{}_j}{\partial u^A}$(id.). $\xi^i{}_A = a^i{}_{Aj}\, x^j$, and so $\xi^i{}_A(\bar{x}^j) = a^i{}_{Ak} A^k{}_j x^j$ which can be written $a^i{}_{Aj} \dfrac{\partial \bar{x}^k}{\partial x^j} x^j$. Using (5.9) we obtain

$$a^i{}_{Aj} \frac{\partial \bar{x}^j}{\partial x^k} = S_A{}^B \frac{\partial \bar{x}^i}{\partial x^i} a^j{}_{Bk}, \tag{5.12}$$

and this can in general be solved to give

$$S_A{}^B = S_A{}^B\left(a^i{}_{Aj},\ \frac{\partial \bar{x}^i}{\partial x^j},\ \frac{\partial x^i}{\partial \bar{x}^j}\right), \tag{5.13}$$

which can then be regarded as a formula holding for any transformation $x^i \to \bar{x}^i$.

The invariance of $\xi^i{}_A$ under the class of transformations which it itself generates is the same sort of naturality concept as is the invariance of $\omega^{-1} \circ d$ under the images of functions under itself. Analogous to fig. 1 is fig. 2 below:

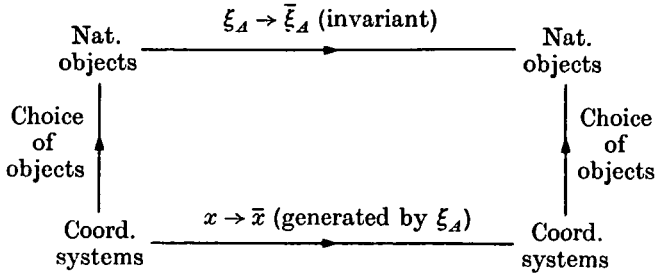


Fig. 2

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

A sub-collection of the vectors $\xi^i{}_A$ which span a subgroup form a natural object with respect to that subgroup; if the subgroup is normal, the corresponding 'normal object' is natural with respect to the whole group. For if $\xi^i{}_A$ is the object and ξ^i an arbitrary member of the larger group, then from (5·9)

$$\pounds_\xi \xi^i{}_A = \xi^i{}_{,j}\,\xi^j{}_A - \xi^i{}_{A,j}\,\xi^j - c_A{}^B \xi^i{}_B \tag{5.14}$$

for some set of constants $c_A{}^B$. So if

$$\pounds_\xi \xi_A = 0 \quad \text{then} \quad [\xi_A, \xi_B] = c_A{}^B \xi_B, \tag{5·15}$$

and this is the condition that ξ_A span a normal subgroup.

As an example, in a Galilean or Lorentz invariant theory, if inertial observer A sets up the object composed of his three, or four, translations (and, if given the usual variational principle, his momentum or energy-momentum object), and inertial observer B does the same with his own three or four translations, then the transformation from the frame of observer A to that of observer B transforms the object of A to the object of B. The same is not true of the rotation object, because the rotations are not a normal subgroup of the invariance group.

Suppose then that, given a set of infinitesimal invariances which make up a natural object, the constants of the motion which are mapped into them by $\omega^{-1} \circ d$ make up a conserved quantity with the same transformation law. Calculating in components

$$\bar{\omega}^{\alpha\beta} \bar{f}_{A,B} = \frac{\partial \bar{x}^\alpha}{\partial x^\gamma} S_A{}^B \omega^{\gamma\delta} \frac{\partial f_B}{\partial x^\delta} + \frac{\partial \bar{x}^\alpha}{\partial x^\gamma} \omega^{\gamma\delta} f_B \frac{\partial S_A{}^B}{\partial x^\delta}. \tag{5·16}$$

Separating into (x^i, y^i) coordinates, and remembering that by (3·11)

$$\bar{x}^i = \bar{x}^i(x^j), \quad \bar{y}^i = \frac{\partial \bar{x}^i}{\partial x^j} y^j$$

$\bar{\omega}^{\alpha\beta} \bar{f}_{A,\beta}$ becomes, after some manipulation

$$\begin{aligned}\Bigg[&\frac{\partial \bar{x}^i}{\partial x^a} S_A{}^B \xi^a{}_B + \frac{\partial \bar{x}^i}{\partial x^a} \omega^{x^a x^b} \frac{\partial S_A{}^B}{\partial x^b} f_B, \\ &\frac{\partial}{\partial \bar{x}^k}\left(\frac{\partial \bar{x}^i}{\partial x^a} S_A{}^B \xi^a{}_B\right) \bar{y}^k + \omega^{y^a x^b} \frac{\partial \bar{x}^i}{\partial x^a} \frac{\partial}{\partial y^b}\left(\frac{\partial S_A{}^B}{\partial x^k} f_B\, y^k\right) \\ &\quad - \omega^{x^a x^b} \frac{\partial \bar{x}^i}{\partial x^a} \frac{\partial}{\partial x^b}\left(\frac{\partial S_A{}^B}{\partial x^k} f_B\, y^k\right) + \omega^{x^a x^b} \frac{\partial}{\partial x^k}\left(\frac{\partial \bar{x}^i}{\partial x^a} \frac{\partial S_A{}^B}{\partial x^b}\right) f_B\, y^k \Bigg].\end{aligned} \tag{5·17}$$

Now given a conserved quantity f_A, it will not be true in general that the transformed $\bar{f}_A$ will remain a conserved quantity for the transformed motion. The condition that it will do so is

$$\begin{aligned}\bar{f}_{A,} \bar{M}_\beta{}^\beta &= \frac{\partial}{\partial \bar{x}^\alpha}(S_A{}^B f_B)\, \bar{M}^\alpha = \frac{\partial}{\partial x^\alpha}(S_A{}^B f_B)\, M^\alpha \\ &= S_A{}^B f_{B,\alpha}\, M^\alpha + \frac{\partial S_A{}^B}{\partial x^\alpha}{}_B M^\alpha.\end{aligned}$$

So if $f_{A,\alpha} M^\alpha = \text{constant}$, then $\bar{f}_{A,\alpha} \bar{M}^\alpha = \text{constant}$, provided

$$\frac{\partial S_A{}^B}{\partial x^\alpha} f_B\, M^\alpha = \text{constant}.$$

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

Since $S_A{}^B = S_A{}^B(x^k)$ and $M^\alpha = (y^i, \eta^i(x^j, y^j))$, this is

$$\frac{\partial S_A{}^B}{\partial x^k} f_B\, y^k = \text{constant}. \tag{5·18}$$

For a non-trivial example consider the (contravariant) angular momentum in 3-space

$$M^{rs} = x^r y^s - x^s y^r.$$

For this $\dfrac{\partial S_A{}^B}{\partial x^k} f_B\, y^k$ corresponds to

$$\frac{\partial}{\partial x^k}\left[\frac{\partial x^r}{\partial \bar{x}^m}\frac{\partial \bar{x}^s}{\partial x^n} - \frac{\partial \bar{x}^r}{\partial x^n}\frac{\partial \bar{x}^s}{\partial x^m}\right] x^m y^n y^k.$$

It may be shown by calculation that if $\bar{x}^r(x^s)$ is homogeneous of degree zero, for each r, so that by Euler's theorem $(\partial \bar{x}^r/\partial x^m)\, x^m = 0$, then the coefficient of $y^n y^k$ in the above, viz.

$$\frac{\partial}{\partial x^k}\left[\frac{\partial \bar{x}^r}{\partial x^m}\frac{\partial \bar{x}^s}{\partial x^n} - \frac{\partial \bar{x}^r}{\partial x^n}\frac{\partial \bar{x}^s}{\partial x^m}\right]$$

is skew-symmetric in (n, k). Hence this is a point transformation for which $(\partial S_A{}^B/\partial x^k) f_B\, y^k = 0$. (Note that this does not imply that $(\partial S_A{}^B/\partial x^k) f_B = 0$.)

Substituting (5·18) into (5·17) we have that for any point transformation for which f_A remains a conserved quantity

$$\bar{\omega}^{\alpha\beta}\bar{f}_{A,\beta} = \left(\frac{\partial \bar{x}^i}{\partial x^a} S_A{}^B\, \xi^a{}_B + \omega^{x^a x^b}\frac{\partial \bar{x}^i}{\partial x^a}\frac{\partial S_A{}^B}{\partial x^k} f_B,\right.$$

$$\left.\frac{\partial}{\partial \bar{x}^k}\left(\frac{\partial \bar{x}^i}{\partial x^a} S_A{}^B\, \xi^a{}_B\right)\bar{y}^k + \omega^{x^a x^b}\frac{\partial}{\partial x^k}\left(\frac{\partial \bar{x}^i}{\partial x^a}\frac{\partial S_A{}^B}{\partial x^b}\right) f_B y^k,\right.$$

and so if f_A corresponds to $((\xi^i{}_A, (\partial \xi^i{}_A/\partial x^k)\, y^k), \bar{f}_A$ corresponds to $(\bar{\xi}^i{}_A, (\partial \bar{\xi}^i{}_A/\partial \bar{x}^k)\bar{y}^k)$ just when $\omega^{x^a x^b} = 0$; and $\omega^{x^a x^b} = 0$ (5·1) are the curl-free conditions that M be derivable from a Lagrangian.

6. *The variational principle and natural transformations*

We have seen that it is always possible to set up a transformation from the constants of the motion to the infinitesimal invariances (Section 3, part iii); more, it is always possible to set up a transformation of the type $\omega^{-1} \circ d$, where ω is a closed 2-form (beginning of Section 4). The significance which was attached to the transformation being given by a closed 2-form is that it is then itself invariant under the same invariances (Section 3, part i). Let us now demonstrate the converse, that if $\omega^{-1} \circ d$ is a linear isomorphism having this naturality property, then ω is a closed 2-form.

Let f, g, etc., be the (up-to-a-factor) constant of the motion, F, G, etc. the corresponding vector fields. Continual use will be made of the formula for the Lie derivative of a form, and of the Leibnitz rule for the Lie derivative:

$$\pounds_F \omega = d(F \rfloor \omega) + F \rfloor d\omega \tag{3·3}$$

$$\pounds_F(R \rfloor S) = (\pounds_F R) \rfloor S + R \rfloor (\pounds_F S) \tag{6·1}$$

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

Fig. 3 below summarizes the naturality:

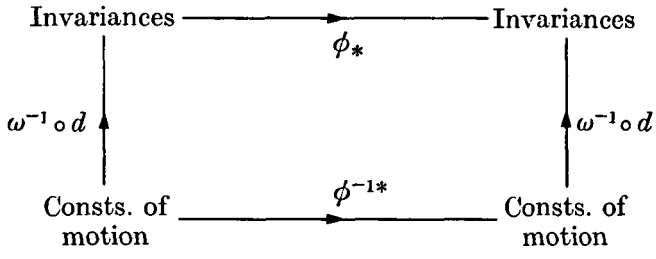


Fig. 3

PROPOSITION 1. *$\omega^{-1} \circ d$ preserves Lie derivatives with respect to any $F \in im(\omega^{-1} \circ d)$.*

Proof. Let F_s^{-1*}, F_{*s} be the 1-parameter groups of induced maps generated by F. By the commutation diagram (Fig. 3)

$$(F_{*s} G) \rfloor \omega = dF_s^{*-1} g = F_s^{*-1} dg.$$

Take $(d/ds)|_{s=0}$ of both sides:

$$(\pounds_F G) \rfloor \omega = d\pounds_F g. \blacksquare$$

PROPOSITION 2. *$\pounds_F \omega = 0$, for all $F \in im(\omega^{-1} \circ d)$.*

Proof. $d\pounds_F g = \pounds_F dg = \pounds_F (G \rfloor \omega) = (\pounds_F G) \rfloor \omega + G \rfloor \pounds_F \omega$, since $\pounds$ obeys the Leibnitz law (6·1). Then from Proposition 1, $G \rfloor \pounds_F \omega = 0$, for all $G \in im(\omega^{-1} \circ d)$. Therefore

$$\pounds_F \omega = 0. \blacksquare$$

We can now show that such an ω must be a closed 2-form.

LEMMA. *If f, g, are constants of the motion, so is $\pounds_F g$.*

Proof. From the formula $\pounds_M \pounds_F g = \pounds_F \pounds_M g + \pounds_{[F, M]} g$.

PROPOSITION 3. *$im(\omega^{-1} \circ d)$ is closed under the Lie bracket.*

Proof. $[F, G] = \pounds_F G = (\omega^{-1} \circ d)(\pounds_F g)$, and $\pounds_F g$ is a constant of the motion, by the Lemma.

Define the Poisson bracket $[f, g]$ by $[f, g] = \pounds_F g$; or equivalently by

$$\omega^{-1} \circ d\,[f, g] = [F, G] \tag{6·2}$$

PROPOSITION 4. *The Poisson bracket is skew-symmetric, and satisfies the Jacobi identities, provided $\pounds_F f = 0$ for at least one $f \in im(\omega^{-1} \circ d)$.*

Proof. By imitation of the properties of the Lie bracket,

$$d[f, g] = -d[g, f]$$

and so $[f, g] = -[g, f] + \text{constant}$; and similarly for the Jacobi identities. The constant $= 0$ by the extra condition. $\blacksquare$

(This extra condition must be taken as a slight strengthening of the assumptions.)

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

PROPOSITION 5. *ω is a closed 2-form.*

Proof. By the properties of the Poisson bracket:

$$[f,g] = \omega^{-1}(df,dg) \quad \text{or} \quad \omega^{\alpha\beta} f_{,\alpha}\, g_{,\beta} = -\omega^{-1}(dg,df),$$

and so $\omega_{\alpha\beta}$ is skew-symmetric. $d\omega = 0$ follows from the Jacobi identities; this proof is best done with ω^{-1}, and in the index notation. An alternative proof in modern notation (suggested to the author by F. A. E. Pirani) is as follows:

$$\pounds_F\, \omega = 0 \text{ (Proposition 2)} \quad \text{and} \quad d(F \rfloor \omega) = d^2 f = 0.$$

Then from $\pounds_F\, \omega = d(F \rfloor \omega) + F \rfloor d\omega$, we have $F \rfloor d\omega = 0$, for all $F \in im\,(\omega^{-1} \circ d)$. Therefore $d\omega = 0$. ∎

The next part is to show that $\omega^{-1} \circ d$, which has so far been defined on the up-to-a-factor constants of the motion, can be extended to a map from all functions on $T(U)$ into a larger class of vector fields, with the constants corresponding to invariances, thus recovering usual Hamiltonian theory.

The constants make up a functionally complete set, and their gradients span the 1-forms at each point. Thus ω^{-1} can certainly be extended to a map from all 1-forms into the vector fields, and it is still a linear isomorphism. It remains to be shown that the naturality continues to hold for a transformation generated by any canonical F (rather than just a canonical invariance).

PROPOSITION 6. *Let f be an arbitrary function, $F = \omega^{-1} \circ df$, and g, G, a constant and its corresponding invariance. Then*

$$\pounds_F\, G = \omega^{-1} \circ d\,(\pounds_F\, g).$$

Proof. $\pounds_F\, G = [F,G] = -[G,F] = -\pounds_G\, F.$
Therefore $(\pounds_F\, G) \rfloor \omega = -(\pounds_G\, F) \rfloor \omega = -\pounds_G (F \rfloor \omega)$, since $\pounds_G \omega = 0$,

$$= -\pounds_G\, df = -d[g,f] = d[f,g] = \pounds_F\, g. \text{ ∎}$$

Propositions 1–5 how hold for all functions and their associated vector fields.

Thus we have the desired result that the existence of the natural transformation (with the motion M itself one of the image vector fields) is equivalent to the existence of a closed 2-form and a function, ω and H, such that $M \rfloor \omega = dH$ (2·5). It is always possible to introduce such a structure, as shown at the beginning of Section 4: the method was to introduce a set of comoving coordinates h^α, and create in these a 1-form θ so that $\pounds_M\, \theta = 0$.

PROPOSITION 7. Any 1-form θ such that $\pounds_M\, \theta = 0$ may be written

$$\theta = \theta_\alpha(h^2, h^3, \ldots, h^{2n})\, dh^\alpha,$$

where h^α are independent up-to-a-factor constants of the motion.

Proof. If θ is of the above form, then $\pounds_M\, \theta = 0$. Conversely, if

$$0 = \pounds_M(\theta_\alpha\, dh^\alpha) = (\pounds_M\, \theta_\alpha)\, dh^\alpha + \theta_\alpha\, d\pounds_M\, h^\alpha = (\pounds_M\, \theta_\alpha)\, dh^\alpha;$$

the result follows since dh^α are linearly independent at each point.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

Since a choice of θ and its associated ω in comoving coordinates is universal for all dynamical systems, to show that one can be chosen with $\det \omega \neq 0$, it is only necessary to take a valid example for each degree of freedom.

$M \rfloor \omega = dH$ now follows, since $\pounds_M \omega = 0$ implies $d(M \rfloor \omega) = 0$ *by* (3·3), and so $M \rfloor \omega = dH$ for some H.

It ceases to be the case that a transformation of the desired type can always be set up as soon as one requires, in addition to the self-invariance property, characterized by a closed 2-form, the curl-free conditions $\omega_{y^i y^j} = 0$ (4.10). Recall the significance which has now been attached to the curl-free conditions: given the fact that out of the infinitesimal point invariances one can construct certain geometric objects (under point transformations), if the curl-free conditions are satisfied, the associated constants of the motion form geometric objects of corresponding type (at least, under transformations for which the transformed concomitant of constants of the motion remains a concomitant of constants).

If there is a natural isomorphism which preserves these geometric objects, then equivalently there exists a closed 2-form ω and a function H, with $M \rfloor \omega = dH$ (2·5) and $\omega_{y^i y^j} = 0$ (4·10). From (2·5) and (4·10) follows the existence of a Lagrangian. This was established by the geometrical argument, using Darboux's theorem, of Section 4, in which the curl-free conditions were obtained. Alternatively, once given (2·5) and (4·10), one may demonstrate the existence of the Lagrangian by direct calculation, as outlined below.

THEOREM. *$M \rfloor \omega = dH$, with $M = (y^i, \eta^i)$, and ω a non-singular closed 2-form for which $\omega_{y^i y^j} = 0$, implies that there exists a function L such that M is given by the Euler–Langrange equations of L.*

Sketch proof.

(i) $M \rfloor \omega = dH$ implies $\pounds_M \omega = 0$.

(ii) $(\pounds_M \omega)_{y^i y^j} = 0$ implies $\omega_{x^i y^j} = \omega_{x^j y^i}$.

(iii) $(d\omega)_{x^i y^j y^k} = 0$ and $\omega_{y^i y^j} = 0$ together imply $\omega_{x^i y^j, y^k} - \omega_{x^i y^k, y^j} = 0$, and therefore $\omega_{x^i y^j} = \dfrac{\partial p^i}{\partial y^j}$ for some p_i.

(iv) From (ii), $p_i = -\dfrac{\partial L}{\partial y^i}$ for some function L.

(v) $(d\omega)_{x^i x^j y^k} = 0$ gives $\omega_{x^i x^j} = \dfrac{\partial^2 L}{\partial x^i \partial y^j} - \dfrac{\partial^2 L}{\partial y^i \partial x^j} + c_{ij}(x^k)$, and from $(d\omega)_{x^i x^j x^k} = 0$ we have $c_{ij} = \dfrac{\partial c_i}{\partial x^j} - \dfrac{\partial c_j}{\partial x^i}$.

(vi) Thus

$$\omega_{\alpha\beta} = \begin{pmatrix} \dfrac{\partial^2 L}{\partial x^i \partial y^j} - \dfrac{\partial^2 L}{\partial y^i \partial x^j} & -\dfrac{\partial^2 L}{\partial y^i \partial y^j} \\ \dfrac{\partial^2 L}{\partial y^i \partial y^j} & 0 \end{pmatrix}$$

with the transformation $L \to L + c_i(x^j)\, y^i$, and this is of the form (2·12).

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

(vii) Substitute for ω into $M \rfloor \omega = dH$ to obtain $H = (\partial L/\partial y^i)\, y^i - L + \gamma(x^j)$ from one half of the equations.

(viii) Substitute for H into the other half of $\mathrm{M} \rfloor \omega = dH$, and the result is the Euler–Lagrange equations in $\bar{L} = L + \gamma_{,i}\, y^i$.

COROLLARY. *Lagrangians are associated with the same 2-form if and only if they differ by a total time derivative plus a constant.*

The proof is either directly from the form of ω above, or by comparing the 1-forms for L and $\bar{L}$:

$$\left(-\frac{\partial L}{\partial y^i}, \quad 0\right) = \left(-\frac{\partial \bar{L}}{\partial y^i}, \quad 0\right) + \left(\frac{\partial \gamma}{\partial x^i}, \quad \frac{\partial \gamma}{\partial y^i}\right).$$

7. *The variational principle and physical theories*

The Variational Principle was once famous for its interpretation as a principle of least (or at any rate, stationary) action. This interpretation has somewhat fallen into disrepute, and has been replaced this century by, if anything, an automatic presumption that any reasonable system of differential equations can be derived from a variational principle. From Section 4 we see that such a belief is untrue, and that such derivability is an unusual, not a general, feature of a system of equations. Yet it is also the case that systems proposed for physical theories always seem to be derivable from a variational principle. Moreover, it is always the 'same' variational principle – the one which associates time translation with energy, space translation with momentum, and so on. Can the foregoing theory throw any light on this strange fact? While it is obviously necessary to extend the work past the limited scope of the present paper, at least as far as field theory, the author feels it is already possible to make a tentative suggestion on this matter.

Let us imagine the setting up of equations for some classical theory. Presumably the end result will have to satisfy the following criteria.

(1) It will admit whatever is considered to be the correct symmetry group of point transformations (e.g. the Galilean or Lorentz group). This creates by the theory of Section 5 various 'natural' objects, such as the translation object.

(2) The classical conserved quantities, with the expected geometric structure, are built into it. (The equations are no doubt intended to express, e.g. rate of change of momentum vector = external force.)

Thus the assumption is made that the invariance objects and the conserved quantities exist as objects of corresponding type. As we saw in section 3 part (iii), it is always possible to set up *some* association between infinitesimal invariances and constants of the motion, but now we are choosing a more particular one.

It seems reasonable to add

(3) The association is, in some sense, the same for all inertial observers. This is the self-invariance property of Section 3 part (i).

According to the preceding work, these assumptions are equivalent to the existence of a variational principle.

The geometric structure of the invariance objects follows purely by group theory

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press

from the structure of the invariance group. The structure of the standard conserved quantities is apparently unconnected (until one assumes the variational principle). It is hard to analyse just why we consider the standard conserved quantities to have the geometric structure that they do, but if the reader will allow that there is at least an element of the empirical in it, then it follows that it is a physical fact about the world – not a purely logical or mathematical fact which would be true in any thinkable universe, but a physical fact which happens to be true in this one – that the standard conserved quantities exist as geometric objects of just the same type as the natural objects set up by the invariance group of the world.

But it is surely an unbelievable coincidence that the two should just happen to be the same. It must be that there is an objective relationship between symmetries and conservation laws. And the variational principle, as a transformation principle between them, is not just a mathematical curiosity, but is the expression of a physical law.

I wish to thank Professors F. A. E. Pirani and C. W. Kilmister for continual help and encouragement; and Dr D. C. Robinson, Dr P. J. McCarthy, Dr C. Bushnell and Mr Philip Maher for helpful conversations.

REFERENCES

(1) LANG, S. *Differentiable Manifolds*. Addison Wesley, Reading Mass., 1972, (p. 134).
(2) SANTILLI, R. M. *Foundations of theoretical mechanics I: The Inverse Problem in Newtonian Mechanics*. Springer-Verlag, New York, 1978.

https://doi.org/10.1017/S0305004100058990 Published online by Cambridge University Press