A TECHNIQUE FOR THE ANALYSIS OF THE INVARIANCE IDENTITIES OF CLASSICAL GAUGE FIELD THEORY BY MEANS OF FUNCTIONAL EQUATIONS

by

David Paul Stapleton

A Dissertation Submitted to the Faculty of the

COMMITTEE ON APPLIED MATHEMATICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1990

Order Number 9103052

A technique for the analysis of the invariance identities of classical gauge field theory by means of functional equations

Stapleton, David Paul, Ph.D.

The University of Arizona, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A TECHNIQUE FOR THE ANALYSIS OF THE INVARIANCE IDENTITIES OF CLASSICAL GAUGE FIELD THEORY BY MEANS OF FUNCTIONAL EQUATIONS

by

David Paul Stapleton

A Dissertation Submitted to the Faculty of the

COMMITTEE ON APPLIED MATHEMATICS

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1990

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by David P. Stapleton

entitled A TECHNIQUE FOR THE ANALYSIS OF THE INVARIANCE IDENTITIES OF CLASSICAL GAUGE FIELD THEORY BY MEANS OF FUNCTIONAL EQUATIONS

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Hanno Rund — Date 8/2/90

W.M. Greenlee — Date 8/2/90

Oma Hamara — Date 8/2/90

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director Hanno Rund — Date 8/2/90

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgement the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: David P. Stapleton

ACKNOWLEDGEMENTS

I wish to thank my advisor, Prof. Hanno Rund, for his apt instructions and for his great patience in reviewing my work. This effort would have been impossible without him. I also owe a great deal to my wife Debby, whose love and support during the stressful time of finishing this document has been invaluable.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

ABSTRACT

In order to obtain the equations of motion for a particle in a classical gauge field, a variational principle is considered. The theory is general in that the structural group is an arbitrary r-dimensional Lie group and the base space is an arbitrary n-dimensional pseudo-Riemannian manifold. An $n+r$ dimensional principal fiber bundle is constructed in order to introduce the usual gauge potentials and field strengths. In addition, a set of r quantities (called "coupling parameters") which transform as the components of an adjoint type (0,1) object and also depend upon the parameter of the particle's trajectory are constructed. The gauge potentials and coupling parameters are evaluated on the identity section of the principal bundle, and the Lagrangian is assumed to be a C^3 scalar function of these and of the components of the metric tensor and tangent vector on the base space. The Lagrangian is not gauge-invariant, but it is stipulated that when the arguments of the Euler-Lagrange vector (evaluated on the identity section) are replaced by their counterparts (which may be evaluated on an arbitrary section) the resulting vector must be gauge-invariant.

A novel application of methods from the theory of functional equations is applied together with standard techniques inherent in the theory of differential equations to show that the arguments of the Lagrangian must occur together in certain prescribed combinations. The invariance postulates uniquely determine the Lagrangian in terms of its arguments other than the coupling parameters and r functions of the coupling parameters. The Lagrangian is shown to separate into a free-field term and an interaction term, and the functions of the coupling parameters are found to be the components of an adjoint type (0,1) quantity whose adjoint absolute derivative vanishes. This agrees with the equations

of certain approaches to the Yang-Mills theory for isotopic spin particles.[1] Standard initial conditions are shown to determine a unique (local) solution to the derived equations of motion .

1 The equations have the same formal structure as systems obtained in the classical limit of quantum mechanical results found by Wong [1], pp. 691-693.

CHAPTER 0

THE ELEMENTS OF CLASSICAL GAUGE FIELD THEORY

This chapter provides a review of the basic aspects of classical gauge field theory. Since its purpose is to introduce notation and to provide a collection of known results for later reference, proofs are omitted.

Our review has the following outline. We begin by introducing an n-dimensional pseudo-Riemannian manifold M (called the "base manifold") and an r-dimensional Lie group G (called the "structural group"). Next, the adjoint representation of G is presented. The so-called "principal fiber bundle" B, corresponding to M and G is introduced, and on it a set of 1-form fields and a set of related 2-form fields are constructed. These forms induce on M corresponding sets of 1-forms and 2-forms, whose components are the gauge potentials and field strengths respectively. The concept of a gauge transformation as an induced transformation of the gauge potentials when G acts on itself by left multiplication (left translation) is put forth.

Except where otherwise stated, the summation convention is employed.

0.1 The Base Manifold

We begin by considering an arbitrary pseudo-Riemannian manifold M, whose dimension n is greater than one.[2] This manifold is referred to as the *base manifold*. We

[2] This section is essentially a review of tensors on pseudo-Riemannian manifolds. The reader is

denote a chart from an atlas on M by a pair (U,h) in which U refers to an open set and h is a homeomorphism $h : U \to R^n$. Lower case Latin indices are used to represent arbitrary numbers from 1 through n, and we denote the coordinates of arbitrary point $x \in U$ by x^i.

A *coordinate transformation* on U is given by

$$\boxed{\bar{x}^k = \bar{x}^k(x^i)}$$

Such a transformation is said to be *admissible* if the functions $\bar{x}^i$ are C^∞ and the Jacobian of the transformation does not vanish on U. In this case, the inverse of the coordinate transformation is denoted by

$$x^i = x^i(\bar{x}^k)$$

and we define the functions

$$\boxed{B_i^k = \frac{\partial x^k}{\partial \bar{x}^i} \quad ; \quad \hat{B}_j^i = \frac{\partial \bar{x}^i}{\partial x^j}} \tag{0.1}$$

Since the matrices (0.1) are inverses, i.e.

$$\hat{B}_m^h B_q^m = \delta_q^h \tag{0.2}$$

it follows (by differentiation of (0.2) with respect to B_j^i) that

$$\boxed{\frac{\partial \hat{B}_m^h}{\partial B_j^i} = -\hat{B}_i^h \hat{B}_m^j} \tag{0.3}$$

We also choose the notation

---

referred to any text on tensors (e.g. Lovelock and Rund [2]) for more details.

$$\boxed{B_{j\,k}^{\,i} = \frac{\partial^2 x^i}{\partial \bar{x}^j \, \partial \bar{x}^k} \quad , \quad \hat{B}_{j\,k}^{\,i} = \frac{\partial^2 \bar{x}^i}{\partial x^j \, \partial x^k}} \tag{0.4}$$

The components of the metric tensor on M are denoted by

$$g_{ij} = g_{ij}(x^m) \tag{0.5}$$

Since M is a pseudo-Riemannian manifold, these functions are such that

$$\boxed{\bar{g}_{ij} = g_{mk} B_i^m B_j^k} \tag{0.6}$$

$$\boxed{g_{ij} = g_{ji}} \tag{0.7}$$

and

$$\boxed{\det(g_{ij}) \neq 0} \tag{0.8}$$

The components of the *inverse of the metric tensor* are assigned the usual notation

$$g^{im} = g^{im}(x^h) \tag{0.9}$$

and by differentiation of

$$g^{ik} g_{kj} = \delta_j^i$$

and taking into account the symmetry (0.7), there results

$$\boxed{\frac{\partial g^{ih}}{\partial g_{pq}} = -\frac{1}{2}\left(g^{ip} g^{qh} + g^{ph} g^{iq}\right)} \tag{0.10}$$

A C^3 *curve* on M is represented parametrically by

$$\{x : x^i = x^i(\tau) \text{ for } \tau \in [\tau_o, \tau_1]\}$$

where $x^i(\tau)$ is a C^3 function on an interval $[\tau_o, \tau_1]$, and the corresponding components of the *tangent vector* are written as

$$\dot{x}^i = \frac{dx^i}{d\tau} \tag{0.11}$$

A *change of parameter* is a substitution $\tau \longrightarrow t$, where t is a class C^1 function on $[\tau_o, \tau_1]$ such that

$$\frac{dt}{d\tau} > 0 \tag{0.12}$$

A parameter τ is called *proper time* if

$$g_{ij}\frac{dx^i}{d\tau}\frac{dx^j}{d\tau} = 1 \tag{0.13}$$

which determines τ up to a constant. Under a change of coordinates, the components of the tangent vector satisfy

$$\boxed{\dot{\bar{x}}^i = \hat{B}^i_j \dot{x}^j \quad ; \quad \dot{x}^i = B^i_j \dot{\bar{x}}^j} \tag{0.14}$$

The second derivatives are denoted by

$$\ddot{x}^i = \frac{d\dot{x}^i}{d\tau} = \frac{d^2x^i}{d\tau^2} \tag{0.15}$$

and under a change of coordinates on M, these transform according to the rule

$$\boxed{\ddot{\bar{x}}^j = \ddot{x}^h \hat{B}^j_h + \hat{B}_h{}^j{}_k \dot{x}^h \dot{x}^k} \tag{0.16}$$

Partial derivatives with respect to the coordinates of M are indicated by placing a comma before an index. For example, the derivatives of the components of the metric are indicated by the notation

$$\boxed{g_{jk,m} = \frac{\partial g_{jk}}{\partial x^m}} \tag{0.17}$$

We also introduce the usual *Christoffel symbols*:

$$\Gamma^i_{j\,k} = \frac{1}{2} g^{im} \left(g_{mj,k} + g_{km,j} - g_{jk,m} \right) \tag{0.18}$$

Since these constitute the components of a connection on M, the usual *absolute derivative* of the tangent vector is given by

$$\boxed{\frac{D\dot{x}^h}{D\tau} = \ddot{x}^h + \Gamma^h_{ij}\,\dot{x}^i\,\dot{x}^j} \tag{0.19}$$

We also have from Ricci's lemma (which is a consequence of the definition (0.18)) that

$$Dg_{ij} = dg_{ij} - \left(\Gamma^h_{i\,k} g_{hj} + \Gamma^h_{j\,k} g_{ih}\right) dx^k = 0$$

It follows that $g_{ij}\,\dot{x}^i\,\dot{x}^j$ is constant along a given curve, if and only if

$$g_{ij}\,\dot{x}^i \frac{D\dot{x}^j}{D\tau} = 0 \tag{0.20}$$

It should be noted that in relativistic mechanics the base manifold is usually identified as Minkowski space-time[3]. Such an identification is <u>not</u> assumed here, but is only a special case of the theory we consider.

---

3 See for example, Drechsler and Mayer [3], p. 116.

0.2 The Structural Group

An r-dimensional Lie group G, is chosen, with $r \geq 1$.[4] This group is referred to as the *structural group* since it is to be identified as the typical fiber of a principal fiber bundle (see section 0.4). The coordinates of its identity element e are (as usual) taken to be zero. Any other element u of G has its r coordinates denoted by u^α ($\alpha = 1,...,r$), and lower case Greek indices in general refer to integers from 1 through r.

Since G is a Lie group, the group multiplication defines analytic *composition functions* $\phi^\alpha: G \to G$, such that if $u,v,w \in G$ with $uv = w$ then

$$\phi^\alpha(u^\gamma, v^\gamma) = w^\alpha$$

As indicated by this notation, the composition functions are functions of $2r$ real arguments. The relations $ue = u$ and $ev = v$ imply that these functions satisfy

$$\frac{\partial \phi^\alpha(u^\gamma,0)}{\partial u^\beta} = \delta^\alpha_\beta \quad ; \quad \frac{\partial \phi^\alpha(0,v^\gamma)}{\partial v^\beta} = \delta^\alpha_\beta \tag{0.21}$$

(where "0" denotes evaluation at the identity e subsequent to differentiation). But the analytic nature of the composition functions imposes the even stronger condition that

$$\phi^\alpha(u^\sigma,v^\sigma) = u^\alpha + v^\alpha + \frac{\partial^2 \phi^\alpha(0,0)}{\partial u^\beta \partial v^\gamma} u^\beta v^\gamma + \textit{higher order terms} \tag{0.22}$$

Among the most important quantities associated with G are the *structure constants*,

---

4 See for instance Rund [4], pp. 124-138, for a survey of the material on Lie groups which is presented here - in the same notation.

which are defined by

$$\boxed{C_{\beta\gamma}^{\alpha} = \frac{\partial^2\phi^{\alpha}(0,0)}{\partial u^{\beta}\partial v^{\gamma}} - \frac{\partial^2\phi^{\alpha}(0,0)}{\partial v^{\gamma}\partial u^{\beta}}} \tag{0.23}$$

According to the definition, the structure constants are skew-symmetric:

$$\boxed{C_{\beta\gamma}^{\alpha} = -C_{\gamma\beta}^{\alpha}} \tag{0.24}$$

and due to the associativity of the group they satisfy the *Jacobi identity*:[5]

$$\boxed{C_{\eta\varepsilon}^{\alpha} C_{\beta\gamma}^{\varepsilon} + C_{\beta\varepsilon}^{\alpha} C_{\gamma\eta}^{\varepsilon} + C_{\gamma\varepsilon}^{\alpha} C_{\eta\beta}^{\varepsilon} = 0} \tag{0.25}$$

0.3 The Adjoint Representation

For arbitrary $u,v \in G$ let $w = uv$, and let $T_u(G)$, $T_v(G)$ and $T_w(G)$ refer to the tangent spaces of G at u, v and w respectively. The adjoint representation[6] may be developed from consideration of multiplication in G as follows.

Given any $u \in G$, the group is said to *act on itself by left translation* L_u: $G \to G$ according to the rule

$$L_u v = w$$

for each element $v \in G$. The action of G on itself induces a map

5 See Pontryagin [5], pp. 380-381, for the proof.

6 See for example Sagle and Walde [6], for an introduction to the adjoint representation.

$$(L_u)_*: T_v(G) \rightarrow T_w(G)$$

which is such that for any $X \in T_v(G)$

$$\left((L_u)_* X\right)^\alpha = \frac{\partial \phi^\alpha(u^\gamma, v^\gamma)}{\partial v^\beta} X^\beta$$

Similarly for any given element v of G, G is said to act on itself by the *right translation* $R_v: G \rightarrow G$, according to the rule (for arbitrary $u \in G$)

$$R_v u = w$$

This induces a map

$$(R_v)_*: T_u(G) \rightarrow T_w(G)$$

such that if $Y \in T_u(G)$,

$$\left((R_v)_* Y\right)^\alpha = \frac{\partial \phi^\alpha(u^\gamma, v^\gamma)}{\partial u^\beta} Y^\beta$$

Consequently, the map $(R_{u^{-1}})_*(L_u)_*$ is such that for any $X \in T_e(G)$

$$\left((R_{u^{-1}})_*(L_u)_* X\right)^\alpha = \frac{\partial \phi^\alpha(u^\varepsilon, (u^{-1})^\varepsilon)}{\partial u^\beta} \frac{\partial \phi^\beta(u^\varepsilon, 0)}{\partial v^\gamma} X^\gamma \in T_e(G)$$

(where (u^{-1}) denotes the *inverse* of element u). Upon defining the r x r matrices

$$\boxed{\lambda_\beta^\alpha(u^\varepsilon) = \frac{\partial \phi^\alpha(u^\varepsilon, (u^{-1})^\varepsilon)}{\partial u^\beta} \; ; \; \chi_\beta^\alpha(u^\varepsilon) = \frac{\partial \phi^\alpha(u^\varepsilon, 0)}{\partial v^\beta} \; ; \; G_\beta^\alpha(u^\varepsilon) = \lambda_\gamma^\alpha(u^\varepsilon)\, \chi_\beta^\gamma(u^\varepsilon)} \tag{0.26}$$

the map $(R_{u^{-1}})_*(L_u)_*$ (applied to elements of $T_e(G)$) may be expressed as

$$\left((R_{u^{-1}})_*(L_u)_*\, X\right)^\alpha = G^\alpha_\beta(u^\varepsilon)\, X^\beta \tag{0.27}$$

The set of matrices $\{G^\alpha_\beta(u^\gamma) : u \in G\}$ together with the operation of matrix multiplication, are known to comprise a group which is isomorphic to G. Hence these matrices are said to constitute a *faithful representation* of G, and this representation is called the *adjoint representation*. Because of the isomorphism, we have for any u,v and $w = uv$ in G that

$$G^\alpha_\beta(w^\varepsilon) = G^\alpha_\gamma(u^\varepsilon)\, G^\gamma_\beta(v^\varepsilon) \tag{0.28}$$

By putting (0.21) into (0.26) we obtain the useful relations

$$\lambda^\alpha_\beta(0) = \delta^\alpha_\beta \quad ; \quad \chi^\alpha_\beta(0) = \delta^\alpha_\beta \quad ; \quad G^\alpha_\beta(0) = \delta^\alpha_\beta \tag{0.29}$$

Evidently, the matrices (0.26) are nonsingular in a neighborhood of the identity -

$$\boxed{\det(\lambda^\alpha_\beta) \neq 0 \quad ; \quad \det(\chi^\gamma_\beta) \neq 0 \quad ; \quad \det(G^\alpha_\beta) \neq 0} \tag{0.30}$$

and we henceforth denote their inverses (respectively) by

$$\boxed{\hat{\lambda}^\alpha_\beta(u^\varepsilon) \quad ; \quad \hat{\chi}^\alpha_\beta(u^\varepsilon) \quad ; \quad \hat{G}^\alpha_\beta(u^\varepsilon)} \tag{0.31}$$

Objects of adjoint type are defined in terms of the matrices of the adjoint representation. Consider any set of r^{q+s} functions $S^{\alpha_1 \ldots \alpha_q}_{\beta_1 \ldots \beta_s}(u^\sigma)$, and denote their values at $u = e$ by

$$\overset{\circ}{S}{}^{\alpha_1 \ldots \alpha_q}_{\beta_1 \ldots \beta_s} = S^{\alpha_1 \ldots \alpha_q}_{\beta_1 \ldots \beta_s}(0)$$

If these functions are such that

$$S^{\alpha_1 \ldots \alpha_q}_{\beta_1 \ldots \beta_s} = G^{\alpha_1}_{\gamma_1} \ldots G^{\alpha_q}_{\gamma_q} \hat{G}^{\sigma_1}_{\beta_1} \ldots \hat{G}^{\sigma_s}_{\beta_s} \overset{o}{S}{}^{\gamma_1 \ldots \gamma_q}_{\sigma_1 \ldots \sigma_s} \tag{0.32}$$

whenever the same arguments u^α are taken in each function without an overscript "o" then they are said to be the *components of a type (q,s) adjoint* object. The components of adjoint type (1,0) and (0,1) objects for instance, are such that

$$\boxed{S^\alpha = G^\alpha_\beta \overset{o}{S}{}^\beta} \tag{0.33}$$

$$\boxed{S_\alpha = \hat{G}^\beta_\alpha \overset{o}{S}_\beta} \tag{0.34}$$

Also, we will need the following non-trivial results:[7]

$$\boxed{\frac{\partial G^\alpha_\varepsilon}{\partial u^\gamma} = C_{\eta}{}^{\alpha}{}_{\sigma} \lambda^\eta_\gamma G^\sigma_\varepsilon} \tag{0.35}$$

$$\boxed{\frac{\partial \lambda^\alpha_\beta}{\partial u^\gamma} - \frac{\partial \lambda^\alpha_\gamma}{\partial u^\beta} = C_{\eta}{}^{\alpha}{}_{\varepsilon} \lambda^\eta_\gamma \lambda^\varepsilon_\beta} \tag{0.36}$$

Another useful result is obtained by differentiation of $G^\alpha_\sigma \hat{G}^\sigma_\beta = \delta^\alpha_\beta$ with respect to u^μ, and substitution from (0.35), namely

$$\boxed{\frac{\partial \hat{G}^\gamma_\beta}{\partial u^\mu} = - \hat{G}^\gamma_\alpha C_{\varepsilon}{}^{\alpha}{}_{\beta} \lambda^\varepsilon_\mu} \tag{0.37}$$

7 The proof of these results is found in the survey paper of Rund [4], pp. 132-138.

Equations (0.29) and (0.35) show that the generators of the Lie algebra for the adjoint representation must be the structure constants - i.e.

$$\frac{\partial G_{\varepsilon}^{\alpha}(0)}{\partial u^{\gamma}} = C_{\gamma\varepsilon}^{\alpha} \tag{0.38}$$

and relations (0.32) through (0.35) lead (non-trivially) to the standard identities

$$\boxed{C_{\beta\gamma}^{\alpha} = C_{\varepsilon\sigma}^{\mu}\,\widehat{G}_{\mu}^{\alpha}\,G_{\beta}^{\varepsilon}\,G_{\gamma}^{\sigma}} \tag{0.39}$$

which shows that these constants are components of a type (1,2) adjoint object.

0.4 The Principal Fiber Bundle

It is desired that the action of G on itself induce transformations of the components of certain form fields on M. The induced transformations are determined by means of a principal fiber bundle associated with M and G, and are motivated by the theory of electromagnetism - in which the gauge potentials are the components of 1-form fields on M which transform under the action of the Lie group $U(1)$.

A manifold B, called the *bundle space*, is introduced together with a projection $p : B \rightarrow M$, which is onto. For each $x \in M$, this determines the so-called *fiber over* x,

$$Y_x = p^{-1}(x)$$

It is assumed that for each chart (U,h) from some atlas on M, there is a diffeomorphism

$\phi : U \times G \rightarrow p^{-1}(U)$ which is such that

$$p \circ \phi(x,u) = x$$

for each $x \in U$ and $u \in G$. Each map ϕ in turn defines diffeomorphisms $\phi_x : G \rightarrow Y_x$ by the stipulation that for all $u \in G$,

$$\phi_x(u) = \phi(x,u)$$

It is also assumed that there exists a map $T : B \times G \rightarrow B$, which for any $z \in B$ and $u \in G$ is written $(z,u) \rightarrow z \cdot u$, such that for any $v \in G$,

$$\phi(x,uv) = \phi(x,u) \cdot v$$

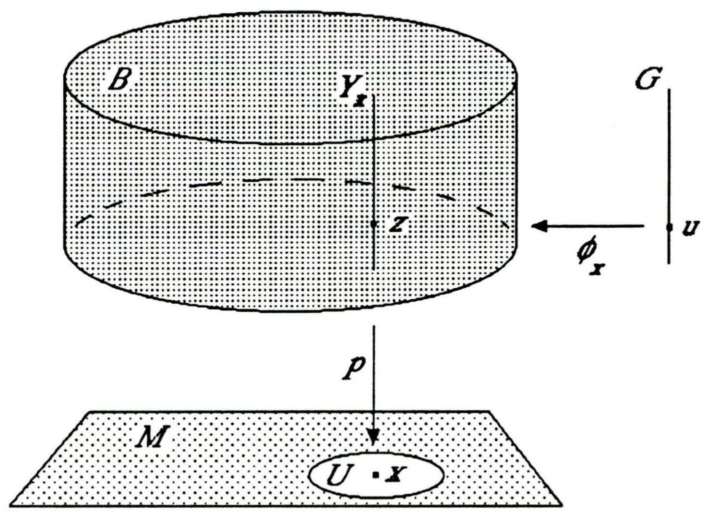
and

$$z \cdot (uv) = (z \cdot u) \cdot v \quad , \quad z \cdot e = z$$

The first stipulation shows that T is a fiber preserving process (i.e. that $p(z \cdot v) = p(z))$ and the last two constitute the requirement that T is a *right action.*

The quadruple (B,M,p,G) together with the map T, is called the *principal fiber bundle* associated with M and G.[8]

[8] The reader is referred to Greub, Halperin and Vanstone [7], for a more complete treatment of principal fiber bundles.



The set of pairs (U,ϕ) (one pair corresponding to each chart on M) is called a *coordinate representation* for the fiber bundle. Corresponding to any chart (V,h') on G such that $u \in V$ and chart (U,h) on M such that $x \in U$, the diffeomorphisms ϕ_x determine that the coordinates of the point $\phi_x(u) \in Y_x$ are C^∞ functions of the coordinates of x on U:

$$\boxed{u^\alpha = u^\alpha(x^h)} \tag{0.40}$$

We note that when $u = e$ these coordinates vanish identically, due to our previous stipulation that the coordinates of the identity element of G vanish. The point $z = \phi(x,u) \in B$ can be described by the $n+r$ coordinates:

$$\boxed{z^J = \begin{cases} x^i & \text{if } J = i \quad (i \in \{1,\ldots,n\}) \\ u^\alpha & \text{if } J = n + \alpha \quad (\alpha \in \{1,\ldots,r\}) \end{cases}} \tag{0.41}$$

We shall frequently require the derivatives of the functions (0.40). Therefore, we introduce the notation

$$\boxed{u_j^\alpha = \frac{\partial u^\alpha}{\partial x^j} \quad , \quad u_{jk}^\alpha = \frac{\partial^2 u^\alpha}{\partial x^j \partial x^k}} \tag{0.42}$$

These derivatives vanish when $u = e$, since the functions (0.40) vanish identically in this case.

A *(global) section* of the fiber bundle is defined to be a map $\sigma: M \rightarrow B$, which is such that for each $x \in M$

$$p \circ \sigma(x) = x$$

Consider also a chart (U,h) on M, and the quadruple (B',U,p',G) where

$$B' = \bigcup_{x \in U} Y_x$$

and $p': B' \rightarrow U$ is the restriction of p to the domain U. This quadruple together with the map $T': B' \times G \rightarrow B'$ which is the restriction of T to the domain $B' \times G$, has the structure of a principal fiber bundle. Hence, for each $u \in G$ there is defined a section σ_u of this local fiber bundle structure, by the requirement that

$$\sigma_u(x) = \phi(x,u) \tag{0.43}$$

whenever $x \in U$ (note that the subscript u denotes an element of G, and is not an index). The *identity section*, σ_e, is therefore the map defined by

$$\sigma_e(x) = \phi(x,e)$$

Since each point in $\sigma_e(U)$ is such that $u^\alpha = 0$, $u^\alpha_h = 0$, $u_h{}^\alpha{}_k = 0$, a superscript "o" is placed over any function of u^α, u^α_h, $u_h{}^\alpha{}_k$ when the function is evaluated at a point on the identity section. For example if $f = f(x^i, u^\alpha, u^\alpha_h, u_h{}^\alpha{}_k)$ then we use the notation

$$\mathring{f} = f(x^i, 0, 0, 0)$$

0.5 The Gauge Fields and the Action of the Structural Group

Through the construction of forms on the manifolds M and B, we can now introduce the gauge fields.[9] Let z represent an arbitrary point of the bundle space B, so that $z = \phi_x{}^{-1}(u)$ for some point x in some chart (U,h) of M and some point u of G. Next, let $\grave{\Lambda}^1$ denote the set of all C^∞ 1-form fields on B, $\grave{\Lambda}^1_u$ the set of C^∞ 1-form fields on $\sigma_u(U)$, and Λ^1 the set of C^∞ 1-form fields on U. It is supposed that we are given a set of C^∞ 1-form fields on B

$$\grave{\omega}^\alpha = G^\alpha_\beta(u^\sigma)\, \mathring{A}^\beta_j(x^i)\, dz^j - \lambda^\alpha_\beta(u^\sigma)\, dz^{n+\beta} \tag{0.44}$$

in which the functions $\mathring{A}^\beta_j$ are the components of an arbitrary type (0,1) vector field on M and G^α_β , λ^α_β are the matrices (0.26).

The gauge fields are uniquely defined by the 1-form fields (0.44) as follows. Each section σ_u induces a map σ_u^*: $\grave{\Lambda}^1_u \to \Lambda^1$ such that any $\grave{\Omega} \in \grave{\Lambda}^1_u \subset \grave{\Lambda}^1$, namely

$$\grave{\Omega} = \grave{\Omega}_i(x^j, u^\sigma)\, dz^i + \grave{\Omega}_{n+\beta}(x^j, u^\sigma)\, dz^{n+\beta}$$

is mapped to

9 See Rund [8] for a complete geometrical development of the gauge potentials.

$$\sigma_u^*(\grave{\Omega}) = \grave{\Omega}_i(x^j, u^\alpha(x^h))dx^i + \grave{\Omega}_{n+\sigma}(x^j, u^\alpha(x^h))\, u_j^\sigma dx^j$$

Thus σ_u^* maps the 1-forms (0.44) to

$$\omega^\alpha = G_\beta^\alpha(u^\sigma(x^i))\mathring{A}_j^\beta(x^i)\, dx^j - \lambda_\beta^\alpha(u^\sigma(x^i))u_j^\beta(x^i)dx^j \qquad (0.45)$$

If we write $\omega^\alpha = A_j^\alpha\, dx^j$, then the components of these 1-form fields on M,

$$A_j^\alpha = A_j^\alpha(x^i, u^\sigma(x^i), u_h^\sigma(x^i))$$

evidently satisfy

$$\boxed{A_j^\alpha = G_\beta^\alpha \mathring{A}_j^\beta - \lambda_\beta^\alpha u_j^\beta} \qquad (0.46)$$

The components A_j^α are called the *gauge fields* (or *gauge potentials*) of our theory, and are completely determined by our choice of the functions $\mathring{A}_j^\beta$ in (0.44) and our choice of section σ_u. Thus, when G acts on itself by left translation ($v \to L_u v = uv = w$) an induced transformation of the gauge potentials occurs:

$$A_j^\alpha(x^j, v^\sigma(x^i), v_j^\sigma(x^i)) \to A_j^\alpha(x^j, w^\sigma(x^i), w_j^\sigma(x^i))$$

This is the called a *(classical) gauge transformation.* Because of the assumption that the components $\mathring{A}_j^\beta$ are tensorial, the definition of σ_u^* implies that under a coordinate transformation on M

$$\boxed{\overline{A}_j^\alpha = B_j^h\, A_h^\alpha} \qquad (0.47)$$

The partial derivatives of the gauge potentials (to be consistent with (0.17)) are denoted by

$$\boxed{A^{\alpha}_{j,k} = \frac{\partial A^{\alpha}_j(x^i,u^{\sigma},u^{\sigma}_q)}{\partial x^k} + \frac{\partial A^{\alpha}_j(x^i,u^{\sigma},u^{\sigma}_q)}{\partial u^{\gamma}} u^{\gamma}_k + \frac{\partial A^{\alpha}_j(x^i,u^{\sigma},u^{\sigma}_q)}{\partial u^{\gamma}_m} u^{\gamma}_{m\,k}} \quad (0.48)$$

and since on the identity section the group parameters and their derivatives vanish, we have

$$\mathring{A}^{\alpha}_{j,k} = \frac{\partial \mathring{A}^{\alpha}_j(x^i)}{\partial x^k}$$

Differentiation of (0.47) shows that

$$\boxed{\overline{A}^{\alpha}_{j,k} = A^{\alpha}_{q,p} B^q_j B^p_k + A^{\alpha}_q B^q_{j\,k}} \quad (0.49)$$

Also, from (0.46) we find that

$$A^{\alpha}_{j,k} = \left(\frac{\partial G^{\alpha}_{\beta}}{\partial u^{\sigma}} u^{\sigma}_k \mathring{A}^{\beta}_j + G^{\alpha}_{\beta} \mathring{A}^{\beta}_{j,k}\right) - \left(\frac{\partial \lambda^{\alpha}_{\beta}}{\partial u^{\sigma}} u^{\sigma}_k u^{\beta}_j + \lambda^{\alpha}_{\beta} u^{\beta}_{j\,k}\right)$$

or, by (0.35)

$$\boxed{A^{\alpha}_{j,k} = \left(G^{\gamma}_{\beta} \mathring{A}^{\beta}_j C_{\mu}{}^{\alpha}_{\gamma} \lambda^{\mu}_{\sigma} u^{\sigma}_k + G^{\alpha}_{\beta} \mathring{A}^{\beta}_{j,k}\right) - \left(\frac{\partial \lambda^{\alpha}_{\beta}}{\partial u^{\sigma}} u^{\sigma}_k u^{\beta}_j + \lambda^{\alpha}_{\beta} u^{\beta}_{j\,k}\right)} \quad (0.50)$$

An important set of quantities related to the gauge potentials is the set of *field strengths* $\{F_j{}^{\alpha}_k\}$. These are defined in terms of the 2-forms

$$F^{\alpha} = d\omega^{\alpha} + \frac{1}{2} C_{\beta}{}^{\alpha}_{\gamma} \omega^{\beta} \wedge \omega^{\gamma}$$

by writing

$$F^{\alpha} = -\frac{1}{2} F_j{}^{\alpha}_k \, dx^j \wedge dx^k$$

so that

$$\boxed{F_{jk}^{\alpha} = A_{j,k}^{\alpha} - A_{k,j}^{\alpha} + C_{\beta\gamma}^{\ \alpha} A_k^{\beta} A_j^{\gamma}} \qquad (0.51)$$

Since they are the components of invariant 2-form fields on M, the field strengths transform as components of type (0,2) tensors -

$$\boxed{\bar{F}_{jk}^{\alpha} = F_{ih}^{\alpha} B_j^i B_k^h} \qquad (0.52)$$

It is also well known (by a short calculation using (0.46) and (0.50)) that the field strengths are components of an adjoint type (1,0) object, i.e.

$$\boxed{F_{jk}^{\alpha} = G_{\beta}^{\alpha} \mathring{F}_{ih}^{\beta}} \qquad (0.53)$$

The gauge fields determine an adjoint connection, via the 1-forms

$$\omega_{\gamma}^{\alpha} = C_{\beta\gamma}^{\ \alpha}\, \omega^{\beta}$$

A short computation using (0.35), (0.38), (0.44) and (0.45) shows that these 1-forms have structure identical with that of connection 1-forms:[10]

$$dG_{\beta}^{\alpha} = G_{\sigma}^{\alpha} \mathring{\omega}_{\beta}^{\sigma} - G_{\beta}^{\eta} \omega_{\eta}^{\alpha}$$

and accordingly, the *(adjoint) absolute differential* of any adjoint type (q,s) quantity is defined. For instance, the absolute differential of an adjoint type (0,1) object S (components S_{α}) has components

[10] See Rund [4], pp. 145-146, for this derivation.

$$DS_\alpha = dS_\alpha - S_\eta\, \omega_\alpha^\eta$$

i.e.

$$\boxed{DS_\alpha = dS_\alpha - S_\eta\, C_{\beta\alpha}^{\eta}\, A_k^\beta\, dx^k} \tag{0.54}$$

CHAPTER 1

THE VARIATIONAL PRINCIPLE

A point particle is now supposed to move along a curve in M, and a variational principle is proposed in order to obtain equations of motion. The particle's motion is assumed to be governed by the Euler-Lagrange equations corresponding to the Lagrangian described in this chapter.

1.1 The Coupling Parameters

In order to put forth the Lagrangian, we introduce a set of geometric quantities called "coupling parameters" which occur as arguments of the Lagrangian. This is accomplished by associating with a particle on M which moves along a curve C with parameter τ, a set of r functions

$$\mathring{e}_\alpha = \mathring{e}_\alpha(\tau)$$

We associate with them the functions, called *coupling parameters*:

$$\boxed{e_\alpha = \widehat{G}_\alpha^\sigma(u^\gamma(x^h))\,\mathring{e}_\sigma(\tau)} \tag{1.1}$$

(an analogy will later be established between these and the so-called "coupling constant" or "charge" of electromagnetic theory[11]). Since they depend only upon the scalars u^α and the

[11] The inclusion of these parameter-dependent, adjoint type (0,1) quantities was suggested by Wong [1] in 1970.

parameter τ, they are scalars under coordinate transformations on M:

$$\bar{e}_\alpha = e_\alpha$$

Although the coupling parameters are not regarded as arbitrary functions, it is assumed that they do not depend upon the values of A_j^α, g_{ij}, or $\dot{x}^i$.

We will require the following expressions for the derivatives of the coupling parameters. According to (1.1), (0.37) and (0.42), the partial derivatives of the coupling parameters with respect to x^k are

$$\begin{aligned} e_{\alpha,k} &= \frac{\partial \hat{G}_\alpha^\beta}{\partial u^\sigma} u_k^\sigma \mathring{e}_\beta \\ &= -\hat{G}_\eta^\beta C_{\varepsilon\alpha}^\eta \lambda_\sigma^\varepsilon u_k^\sigma \mathring{e}_\beta \end{aligned}$$

that is

$$\boxed{e_{\alpha,k} = -e_\eta C_{\varepsilon\alpha}^\eta \lambda_\sigma^\varepsilon u_k^\sigma} \qquad (1.2)$$

We also define

$$\boxed{f_\alpha = \frac{de_\alpha}{d\tau} \quad ; \quad \mathring{f}_\alpha = \frac{d\mathring{e}_\alpha}{d\tau}} \qquad (1.3)$$

which, according to (1.1) and (1.2), satisfy

$$\boxed{f_\alpha = \hat{G}_\alpha^\sigma \mathring{f}_\sigma - e_\gamma C_{\mu\alpha}^\gamma \lambda_\varepsilon^\mu u_h^\varepsilon \dot{x}^h} \qquad (1.4)$$

The absolute derivative of the quantity whose components are the coupling parameters has components (according to (1.3) and (0.54))

$$\boxed{\frac{De_\alpha}{D\tau} = f_\alpha - e_\eta\, C^{\eta}_{\beta\alpha}\, A^{\beta}_{k}\dot{x}^{k}} \tag{1.5}$$

1.2 The Lagrangian

It is required that the Lagrangian be of the form

$$\boxed{\mathring{L} = L(\mathring{e}_\alpha, \mathring{A}^{\alpha}_{j}, g_{ij}, \dot{x}^{j})} \tag{1.6}$$

and furthermore be the evaluation at $u^\alpha = 0$ (that is, on the identity section) of a C^3 real-valued function

$$\boxed{L = L(e_\alpha, A^{\alpha}_{j}, g_{ij}, \dot{x}^{j})} \tag{1.7}$$

It is assumed that the function L is defined on an appropriate domain in the following sense. We restrict our analysis to a simply connected open region W of the tangent bundle $T(M)$ called the *admissible region of the tangent bundle,* which is such that for any point $z \in W$ (with $2n$ coordinates x^h, $\dot{x}^h$)

$$\boxed{g_{ij}\,\dot{x}^{i}\,\dot{x}^{j} \neq 0} \tag{1.8}$$

Let $z' \in W$, and let (U,h) be any chart on $T(M)$ with $z' \in U$. Corresponding to each such pair $(z',(U,h))$ there exist pairs $(W_{z'},\varepsilon_{z'})$ for which:

(i) $W_{z'}$ is a neighborhood of z' such that

$$W_{z'} \subset U \cap W$$

(ii) $\varepsilon_{z'}$ is a positive real number,

(iii) for any real numbers $t^i \in (1-\varepsilon_{z'}, 1+\varepsilon_{z'})$ and point $z \in W_{z'}$ with coordinates x^h, $\dot{x}^h$, the $2n$ numbers

$$x^1 , \ldots , x^n , \frac{\dot{x}^1}{t^1} , \ldots , \frac{\dot{x}^n}{t^n}$$

are the coordinates of a point in $U \cap W$.

This can be accomplished for any given pair $(z',(U,h))$ because W is an open set. We shall say, however, that L is defined on an *admissible domain* if for each such pair $(z',(U,h))$ there exists at least one pair $(W_{z'},\varepsilon_{z'})$ as described above, for which the following holds:

(iv) Let $z \in W_{z'}$, and let C be any curve on M whose parameterization is such that the tangent vector at $x = p(z)$ (where $p : T(M) \rightarrow M$ is the projection) has components $\dot{x}^h$ (the last n coordinates of z). Let $u \in G$ be taken arbitrarily and define the functions (with no summation implied)

$$\widetilde{A}_h^\alpha = t^h A_h^\alpha(x^i , u^\sigma(x^i) , u_k^\sigma(x^i))$$

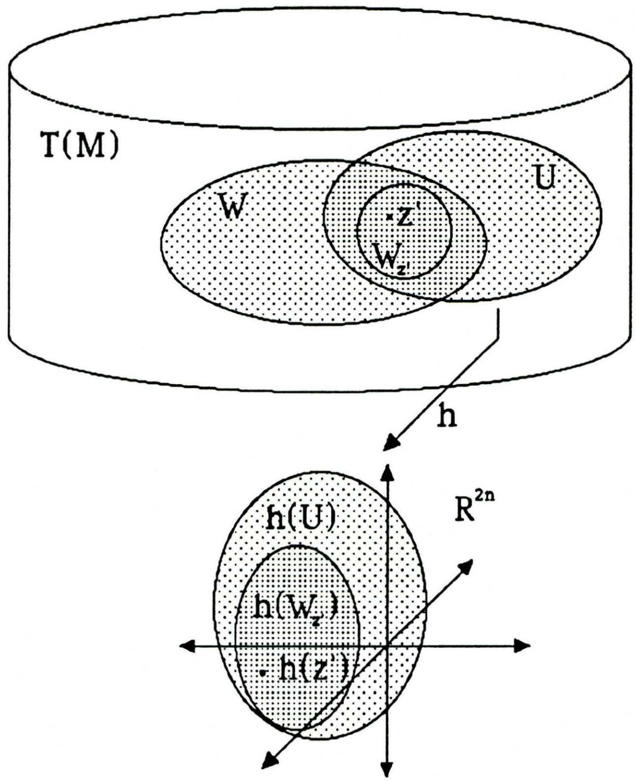
$$\widetilde{g}_{hk} = t^h t^k g_{hk}(x^i)$$

$$\widetilde{\dot{x}}^h = \frac{\dot{x}^h}{t^h}$$

It is required that the function

$$\widetilde{L} = L (e_\alpha , \widetilde{A}_j^\alpha , \widetilde{g}_{ij} , \widetilde{\dot{x}}^j)$$

is defined for all values of the variables $t^i \in (1-\varepsilon_{z'}, 1+\varepsilon_{z'})$.



The function L is restricted by the requirement that for all permissible values of its arguments, there is at least one pair of indices (i,j) $(1 \leq i,j \leq n)$ such that

$$\boxed{\frac{\partial L}{\partial g_{ij}} \neq 0} \tag{1.9}$$

It will be seen in chapter 3 that without this assumption, the usual existence and uniqueness theorem for solutions to the Euler-Lagrange equations cannot be applied.

Furthermore, it is stipulated that the function L must be a scalar under coordinate transformations on M -

$$\boxed{L(e_\alpha, A_j{}^\alpha, g_{ij}, \dot{x}^j) = \overline{L}(e_\alpha, \overline{A}_j{}^\alpha, \overline{g}_{ij}, \dot{\overline{x}}^j)} \tag{1.10}$$

The Lagrangian gives rise to the *fundamental integral*

$$\boxed{I = \int_{\tau_0}^{\tau_1} L(\mathring{e}_\alpha, \mathring{A}_j^\alpha, g_{ij}, \dot{x}^j)\, d\tau} \tag{1.11}$$

and under an admissible change of parameter $\tau \rightarrow t$, this integral is required to be *parameter-invariant* -

$$\boxed{I = \int_{t(\tau_0)}^{t(\tau_1)} L(\mathring{e}_\alpha(t), \mathring{A}_j^\alpha(x^h), g_{ij}(x^h), \dot{x}^j\left(\frac{d\tau}{dt}\right)^{-1})\frac{d\tau}{dt}\, dt}$$

This is known to be possible if and only if[12]

$$\boxed{\frac{\partial \mathring{L}}{\partial \dot{x}^k}\dot{x}^k = \mathring{L}} \tag{1.12}$$

12 By theorem (see e.g. Logan [9], pp. 154-156) the set of equations

$$\frac{\partial L}{\partial \dot{x}^k}\dot{x}^k = L \quad ; \quad \frac{\partial L}{\partial \tau} = 0$$

(known as *Zermelo conditions*) are satisfied by a C^2 function

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

for which the functions x^i are required to be of class C^2, if and only if the fundamental integral

$$I = \int_{\tau_0}^{\tau_1} L(\tau, x^i, \dot{x}^i)\, d\tau$$

is parameter invariant. Since the arguments of our Lagrangian (1.6) can be written entirely in terms of τ, x^i, $\dot{x}^i$ this theorem applies - yielding (1.12) and (1.13).

and also

$$\boxed{\frac{\partial \mathring{L}}{\partial \mathring{e}_\alpha}\mathring{f}_\alpha = 0} \tag{1.13}$$

Finally, the Lagrangian is required to lead to gauge-invariant equations of motion in the following sense. Let

$$\mathring{L} = L(\,\mathring{e}_\alpha(\tau),\, \mathring{A}_j^\alpha(x^h),\, g_{ij}(x^h)\,, \dot{x}^j\,)$$

The *Euler-Lagrange vector* is

$$\mathring{E}_k = \frac{d}{d\tau}\left(\frac{\partial \mathring{L}}{\partial \dot{x}^k}\right) - \frac{\partial \mathring{L}}{\partial x^k} \tag{1.14}$$

Because the derivatives u_h^α vanish on the identity section, (1.2) implies that $\mathring{e}_{\alpha,k} = 0$. Hence (1.14) implies that for some function E_k

$$\mathring{E}_k = E_k(\,\mathring{e}_\alpha\,, \mathring{f}_\alpha\,, \mathring{A}_j^\alpha\,, \mathring{A}_{j,k}^\alpha\,, g_{ij}\,, g_{ij,k}\,, \dot{x}^j\,, \ddot{x}^j\,)$$

When the arguments $\mathring{e}_\alpha\,, \mathring{f}_\alpha\,, \mathring{A}_j^\alpha\,, \mathring{A}_{j,k}^\alpha$ are replaced by $e_\alpha\,, f_\alpha\,, A_j^\alpha\,, A_{j,k}^\alpha$ the resulting vector is denoted by

$$\boxed{E_k = E_k(\,e_\alpha\,, f_\alpha\,, A_j^\alpha\,, A_{j,k}^\alpha\,, g_{ij}\,, g_{ij,k}\,, \dot{x}^j\,, \ddot{x}^j\,)} \tag{1.15}$$

This vector is said to be *gauge-invariant*, if

$$\boxed{E_k = \mathring{E}_k} \tag{1.16}$$

in which case the equations of motion for our particle (the *Euler-Lagrange equations* $\mathring{E}_k = 0$) are equivalent to

$$\boxed{E_k = 0} \tag{1.17}$$

We stipulate therefore that equation (1.16) must hold, and as a result obtain the gauge-invariant form (1.17) for our equations of motion.

Note that the right-hand side of equation (1.16) does not depend upon the choice of $u \in G$, but according to (1.1), (1.4), (0.46) and (0.50) the left-hand side does (through the parameters u^α, u_h^α, $u_{h\,j}^{\,\alpha}$). Differentiation of both sides therefore shows that equations (1.17) are gauge-invariant if and only if

$$\boxed{\frac{\partial E_k}{\partial u^\sigma} = 0} \tag{1.18}$$

$$\boxed{\frac{\partial E_k}{\partial u_h^\sigma} = 0} \tag{1.19}$$

and

$$\boxed{\frac{\partial E_k}{\partial u_{h\,p}^{\,\sigma}} = 0} \tag{1.20}$$

where we have used the notation

$$E_k =$$
$$E_k(e_\alpha(\tau,u^\gamma), f_\alpha(\tau,\dot{x}^i,u^\gamma,u_h^\gamma), A_j^\alpha(x^i,u^\gamma,u_h^\gamma), A_{j,k}^\alpha(x^i,u^\gamma,u_h^\gamma,u_{h\,i}^{\,\gamma}), g_{ij}(x^h), g_{ij,k}(x^h), \dot{x}^j, \ddot{x}^j)$$

It should be noted that in multiple integral gauge field theories, gauge-invariance of the equations of motion is achieved by requiring that the Lagrangian be gauge-invariant. Such an assumption is not made in this paper, however. If the function L of (1.7) were gauge-invariant, we would have (in analogy to (1.19))

$$\frac{\partial L(e_\alpha(\tau,u^\gamma),A_j^\alpha(x^i,u^\gamma,u_h^\gamma),g_{ij}(x^h),\dot{x}^i)}{\partial u_h^\sigma} = 0$$

that is,

$$\frac{\partial L}{\partial A_i^\gamma}\frac{\partial A_i^\gamma}{\partial u_h^\sigma} = 0$$

But, by (0.46), this is

$$-\frac{\partial L}{\partial A_h^\gamma}\lambda_\sigma^\gamma = 0$$

or, by (0.30)

$$\frac{\partial L}{\partial A_h^\gamma} = 0$$

which would indicate that L does not depend on the gauge potentials! We conclude therefore that *gauge-invariant Lagrangians of the form (1.7) do not exist.* Our Lagrangian is therefore taken to be the non-gauge-invariant function (1.6), and we obtain gauge-invariant equations of motion by demanding that the functions (1.15) be gauge-invariant.

1.3 The Euler-Lagrange Vector

For the sake of brevity, we introduce the following notation for the derivatives of L. Let

$$L_k = \frac{\partial L}{\partial \dot{x}^k} \quad ; \quad L_\alpha^k = \frac{\partial L}{\partial A_k^\alpha} \quad ; \quad L_{\alpha m}^k = \frac{\partial L_m}{\partial A_k^\alpha} \quad ; \quad L_{km} = \frac{\partial^2 L}{\partial \dot{x}^k \partial \dot{x}^m}$$

$$\Lambda^\alpha = \frac{\partial L}{\partial e_\alpha} \quad ; \quad \Lambda_k^\alpha = \frac{\partial L_k}{\partial e_\alpha} \quad ; \quad \pounds^{ij} = \frac{\partial L}{\partial g_{ij}} \quad ; \quad \pounds_k^{ij} = \frac{\partial L_k}{\partial g_{ij}} \qquad (1.21)$$

The corresponding derivatives of the Lagrangian are denoted by

$$\mathring{L}_k = \frac{\partial \mathring{L}}{\partial \dot{x}^k} \quad ; \quad \mathring{L}_\alpha^k = \frac{\partial \mathring{L}}{\partial \mathring{A}_k^\alpha} \quad ; \quad \mathring{L}_{\alpha m}^k = \frac{\partial \mathring{L}_m}{\partial \mathring{A}_k^\alpha} \quad ; \quad \mathring{L}_{km} = \frac{\partial^2 \mathring{L}}{\partial \dot{x}^k \partial \dot{x}^m}$$

$$\mathring{\Lambda}^\alpha = \frac{\partial \mathring{L}}{\partial \mathring{e}_\alpha} \quad ; \quad \mathring{\Lambda}_k^\alpha = \frac{\partial \mathring{L}_k}{\partial \mathring{e}_\alpha} \quad ; \quad \mathring{\pounds}^{ij} = \frac{\partial \mathring{L}}{\partial g_{ij}} \quad ; \quad \mathring{\pounds}_k^{ij} = \frac{\partial \mathring{L}_k}{\partial g_{ij}} \qquad (1.22)$$

The Euler-Lagrange expression is therefore calculated as follows. Since the dependence of the Lagrangian on x^k is given by

$$\begin{aligned} \frac{\partial \mathring{L}}{\partial x^k} &= \frac{\partial \mathring{L}}{\partial \mathring{A}_j^\alpha} \mathring{A}_{j,k}^\alpha + \frac{\partial \mathring{L}}{\partial g_{ij}} g_{ij,k} \\ &= \mathring{L}_\alpha^j \mathring{A}_{j,k}^\alpha + \mathring{\pounds}^{ij} g_{ij,k} \end{aligned}$$

while along the trajectory ($x^j = x^j(\tau)$)

$$\begin{aligned} \frac{d\mathring{L}_k}{d\tau} &= \frac{\partial \mathring{L}_k}{\partial \mathring{e}_\alpha} \frac{d\mathring{e}_\alpha}{d\tau} + \frac{\partial \mathring{L}_k}{\partial \mathring{A}_j^\alpha} \mathring{A}_{j,m}^\alpha \dot{x}^m + \frac{\partial \mathring{L}_k}{\partial g_{ij}} g_{ij,m} \dot{x}^m + \frac{\partial \mathring{L}_k}{\partial \dot{x}^m} \ddot{x}^m \\ &= \mathring{\Lambda}_k^\alpha \mathring{f}_\alpha + \mathring{L}_{\alpha k}^j \mathring{A}_{j,m}^\alpha \dot{x}^m + \mathring{\pounds}_k^{ij} g_{ij,m} \dot{x}^m + \mathring{L}_{km} \ddot{x}^m \end{aligned}$$

the Euler-Lagrange vector is

$$\mathring{E}_k = \frac{d\mathring{L}_k}{d\tau} - \frac{\partial \mathring{L}}{\partial x^k}$$

$$= \mathring{\Lambda}_k^{\alpha}\mathring{f}_\alpha + \mathring{L}^j_{\alpha k}\mathring{A}^\alpha_{j,m}\dot{x}^m + \mathring{\pounds}^{ij}_k g_{ij,m}\dot{x}^m + \mathring{L}_{km}\ddot{x}^m - \mathring{L}^j_\alpha \mathring{A}^\alpha_{j,k} - \mathring{\pounds}^{ij} g_{ij,k}$$

i.e.

$$\mathring{E}_k = \mathring{L}_{km}\ddot{x}^m + \mathring{\Lambda}_k^{\alpha}\mathring{f}_\alpha + (\mathring{L}^j_{\alpha k}\dot{x}^m - \mathring{L}^j_\alpha \delta^m_k)\mathring{A}^\alpha_{j,m} + (\mathring{\pounds}^{ij}_k \dot{x}^m - \mathring{\pounds}^{ij}\delta^m_k) g_{ij,m} \tag{1.23}$$

The corresponding vector E_k defined in (1.15) is therefore

$$\boxed{E_k = L_{km}\ddot{x}^m + \Lambda_k^{\alpha} f_\alpha + (L^j_{\alpha k}\dot{x}^m - L^j_\alpha \delta^m_k) A^\alpha_{j,m} + (\pounds^{ij}_k \dot{x}^m - \pounds^{ij}\delta^m_k) g_{ij,m}} \tag{1.24}$$

Unfortunately the quantities with Latin indices in (1.24) are not all components of tensors, and the quantities with Greek indices are not all components of adjoint type objects. A more suitable form of (1.24) will be determined in upcoming chapters.

CHAPTER 2

THE USE OF FUNCTIONAL EQUATIONS TO RESTRICT THE FORM OF THE LAGRANGIAN

In this chapter a new technique is employed, by which it is shown that the function L of (1.7) must be such that

$$L = L^*(e_\alpha , A_j^\alpha \dot{x}^j , g_{ij}\dot{x}^i\dot{x}^j)$$

for some scalar function L^*. Hence, the Lagrangian (1.6) must be of the form

$$\mathring{L} = L^*(\mathring{e}_\alpha , \mathring{A}_j^\alpha \dot{x}^j , g_{ij}\dot{x}^i\dot{x}^j)$$

This derivation requires only the following assumptions from chapter 1:

(i) L is a C^3, real-valued function of the arguments indicated in (1.7),

(ii) L is defined on an admissible domain as indicated in section 1.2 (this restricts us to an open region of the tangent bundle where $g_{ij}\dot{x}^i\dot{x}^j \neq 0$),

(iii) L is a scalar under coordinate transformations on M,

(iv) the vector E_k of (1.15) does not depend on the parameters $u_{h}^{\sigma}{}_{k}$.

As a corollary, it is shown that the additional assumption that:

(v) the fundamental integral (1.11) is parameter-invariant,

implies that $\mathring{L}$ can be expressed as

$$\overset{\circ}{L} = \sqrt{|g_{ij}\dot{x}^i\dot{x}^j|}\,\Lambda(\overset{\circ}{e}_\alpha, \frac{\overset{\circ}{A}{}^\alpha}{\sqrt{|g_{ij}\dot{x}^i\dot{x}^j|}})$$

for some C^3 scalar function Λ.

Of particular importance is the technique used in the derivation of Lemma 1 (section 2.4). By use of this technique, it is shown that the arguments of the function L occur together in certain prescribed combinations. The procedure requires the use of functional equations, and proceeds in a manner which is related to a method for determining how the arguments of homogeneous functions must combine.[13] It demonstrates how one may deduce the combinations of arguments that occur for certain classes of non-gauge-invariant Lagrangians.

2.1 Transformation Laws for the Derivatives of the Lagrangian

In order to shorten calculations, let us begin by obtaining transformation laws for the derivatives of the scalar function L. We recall from (1.10) that

$$L(e_\alpha, A_j{}^\alpha, g_{ij}, \dot{x}^j) = \overline{L}(e_\alpha, \overline{A}_j{}^\alpha, \overline{g}_{ij}, \dot{\overline{x}}^j)$$

and take partial derivatives of both sides with respect to A_p^β, g_{pq} and $\dot{x}^p$ (employing the notation of (0.1) and (1.21)).

Differentiation of (1.7) with respect to the gauge potentials yields (due to (0.47))

[13] The reader is referred to the method for the solution of homogeneous equations of Aczel [10], pp. 229-234.

$$L_\alpha^p = \overline{L}_\sigma^j \frac{\partial(A_k^\sigma B_j^k)}{\partial A_p^\alpha}$$

$$= \overline{L}_\alpha^j B_j^p$$

Next, differentiation with respect to the components of the metric gives (by (0.6) and (0.7))

$$\pounds^{pq} = \overline{\pounds}^{ij} \frac{\partial(\, g_{mk}\, B_i^m\, B_j^k\,)}{\partial g_{pq}}$$

$$= \overline{\pounds}^{ij} \frac{1}{2} (\, \delta_m^p\, \delta_k^q + \delta_k^p\, \delta_m^q\,)\, B_i^m\, B_j^k$$

$$= \overline{\pounds}^{ij}\, B_i^p\, B_j^q$$

Finally, differentiation with respect to the components of the tangent vector shows that (because of (0.14))

$$L_p = \overline{L}_j \frac{\partial(\dot{x}^j\, \widehat{B}_j^k\,)}{\partial \dot{x}^p}$$

$$= \overline{L}_j\, \widehat{B}_p^j$$

In summary, we have obtained the tensorial relations

$$\boxed{L_\beta^p = \overline{L}_\beta^j B_j^p \quad ; \quad \pounds^{pq} = \overline{\pounds}^{ij}\, B_i^p\, B_j^q \quad ; \quad L_p = \overline{L}_j\, \widehat{B}_p^j} \qquad (2.1)$$

The last set of equations in (2.1) is

$$L_i(e_\alpha, A_k^\alpha, g_{km}, \dot{x}^k) = \widehat{B}_i^h \overline{L}_h(e_\alpha, \overline{A}_k^\alpha, \overline{g}_{ij}, \dot{\overline{x}}^j)$$

Upon differentiating this equation with respect to A_p^β, g_{pq} and $\dot{x}^p$ respectively, we obtain the relations

$$L_{\beta i}^p = \widehat{B}_i^h \overline{L}_{\alpha h}^j \frac{\partial(A_k^\alpha B_j^k)}{\partial A_p^\beta}$$

$$= \overline{L}_{\beta h}^j B_j^p \widehat{B}_i^h$$

$$£_i^{pq} = \overline{£}_h^{tj} \widehat{B}_i^h \frac{\partial(g_{mk} B_t^m B_j^k)}{\partial g_{pq}}$$

$$= \overline{£}_h^{tj} \widehat{B}_i^h \frac{1}{2}(\delta_m^p \delta_k^q + \delta_n^p \delta_m^q) B_t^m B_j^k$$

$$= \overline{£}_h^{tj} \widehat{B}_i^h B_t^p B_j^q$$

$$L_{ip} = \overline{L}_{hj} \widehat{B}_i^h \frac{\partial(\dot{x}^k \widehat{B}_k^j)}{\partial \dot{x}^p}$$

$$= \overline{L}_{hj} \widehat{B}_i^h \widehat{B}_p^j$$

Thus, we have deduced the tensorial transformation laws

$$\boxed{L_{\beta i}^p = \overline{L}_{\beta h}^j B_j^p \widehat{B}_i^h \;\; ; \;\; £_i^{pq} = \overline{£}_h^{tj} \widehat{B}_i^h B_t^p B_j^q \;\; ; \;\; L_{ip} = \overline{L}_{hq} \widehat{B}_i^h \widehat{B}_p^q} \qquad (2.2)$$

2.2 Requirements for Scalar Functions

In this section, we obtain a system of partial differential equations whose satisfaction is necessary and sufficient in order that L be a scalar function. The system is obtained by differentiating (1.10) with respect to the independent functions B_q^p of (0.1).

If (1.10) is expressed in the form

$$L\,(\,e_\alpha\,,A_j{}^\alpha,\,g_{ij}\,,\dot{x}^j\,)\;=\;\bar{L}\,(\,e_\alpha\,,A_k{}^\alpha B_j^k,\,g_{mk}B_i^m B_j^k\,,\dot{x}^k\widehat{B}_k^j\,)$$

and differentiated, it is found that

$$0\;=\;\frac{\partial\bar{L}}{\partial\bar{A}_j{}^\alpha}\frac{\partial(A_k^\alpha B_j^k)}{\partial B_q^p}\;+\;\frac{\partial\bar{L}}{\partial\bar{g}_{ij}}\frac{\partial(g_{mk}B_i^m B_j^k)}{\partial B_q^p}\;+\;\frac{\partial\bar{L}}{\partial\bar{\dot{x}}^j}\frac{\partial(\dot{x}^k\widehat{B}_k^j)}{\partial B_q^p}$$

$$=\;\bar{L}_\alpha^j A_k^\alpha\frac{\partial B_j^k}{\partial B_q^p}\;+\;\bar{L}^{ij}g_{mk}\frac{\partial(\,B_i^m B_j^k\,)}{\partial B_q^p}\;+\;\bar{L}_j\dot{x}^k\frac{\partial\widehat{B}_k^j}{\partial B_q^p}$$

Therefore, upon substitution from (0.3)

$$0\;=\;\bar{L}_\alpha^j A_k^\alpha\,\delta_p^k\,\delta_j^q\;+\;\bar{L}^{ij}g_{mk}\,(B_i^m\delta_p^k\,\delta_j^q+B_j^k\,\delta_p^m\,\delta_i^q)\;-\;\bar{L}_j\dot{x}^k\,\widehat{B}_p^j\,\widehat{B}_k^q$$

$$=\;\bar{L}_\alpha^q A_p^\alpha\;+\;\bar{L}^{iq}g_{mp}B_i^m\;+\;\bar{L}^{qj}g_{pk}B_j^k\;-\;L_p\,\dot{x}^k\,\widehat{B}_k^q$$

But due to (2.1) and (2.2) the quantities in this equation are all tensorial, so

$$0\;=\;L_\alpha^m A_p^\alpha\widehat{B}_m^q\;+\;L^{mk}g_{mp}\widehat{B}_k^q\;+\;L^{mk}g_{pk}\widehat{B}_m^q\;-\;L_p\,\dot{x}^m\,\widehat{B}_m^q$$

The middle two terms are identical since g_{pk} and L^{mk} are symmetric, so

$$0 = L^m_\alpha A^\alpha_p \widehat{B}^q_m + 2L^{mk} g_{pk} \widehat{B}^q_m - L_p \dot{x}^m \widehat{B}^q_m$$

and, since for admissible coordinate transformations $\det(\widehat{B}^q_m) \neq 0$, this shows that

$$\boxed{L^m_\alpha A^\alpha_p - L_p \dot{x}^m + 2\pounds^{mk} g_{pk} = 0} \tag{2.3}$$

Evidently, this system of partial differential equations represents a set of necessary and sufficient conditions for the function L to be a scalar. Equations (2.1) and (2.2) indicate that the quantities in this equation are tensorial.

2.3 Euler-Lagrange Vectors Which Satisfy the Third Invariance Identity

We shall now deduce a set of conditions on L, which represent necessary and sufficient conditions for the vector E_k to be independent of the parameters $u^{\alpha}_{h\,k}$ (i.e. satisfy (1.20)).

Substitution of (1.24) into (1.20) gives

$$\frac{\partial}{\partial u^{\sigma}_{h\,p}}\left(L_{km}\ddot{x}^m + \Lambda^{\alpha}_k f_\alpha + (L^j_{\alpha k}\dot{x}^m - L^j_\alpha \delta^m_k)A^{\alpha}_{j,m} + (\pounds^{ij}_k \dot{x}^m - \pounds^{ij}\delta^m_k)g_{ij,m}\right) = 0$$

But according to (1.1), (1.4), (0.46) and (0.50), only the term

$$(L^j_{\alpha k}\dot{x}^m - L^j_\alpha \delta^m_k)A^{\alpha}_{j,m}$$

depends upon $u^{\sigma}_{h\,p}$. Consequently (1.20) holds if and only if

$$(L^j_{\alpha k}\dot{x}^m - L^j_\alpha \delta^m_k)\frac{\partial A^\alpha_{j,m}}{\partial u^\sigma_{h\,p}} = 0 \tag{2.4}$$

or, by (0.50)

$$-\frac{1}{2}(L^j_{\alpha k}\dot{x}^m - L^j_\alpha \delta^m_k)\,\lambda^\alpha_\sigma\left(\delta^j_h\delta^m_p + \delta^j_p\delta^m_h\right) = 0$$

Therefore, due to (0.30), the system

$$\boxed{(L^j_{\alpha k}\dot{x}^m - L^j_\alpha\delta^m_k) + (L^m_{\alpha k}\dot{x}^j - L^m_\alpha\delta^j_k) = 0} \tag{2.5}$$

is a set of necessary and sufficient conditions for the validity of (1.20). The special case $m = j$ of (2.5) yields another useful result (with no summation implied)

$$\boxed{L^j_{\alpha k}\dot{x}^j - L^j_\alpha\delta^j_k = 0} \tag{2.6}$$

2.4 Restricting the Form of the Lagrangian in Special Coordinate Systems

In this section we deduce a lemma, which states that for any point z' in the admissible region W of the tangent bundle (see section 1.2) there exists a neighborhood of z' and a coordinate system, such that the arguments of L occur together in prescribed combinations. We suspend use of the summation convention in obtaining this result. In particular if we write

$$C^\alpha_j = A^\alpha_j\dot{x}^j \ , \ e_{ij} = g_{ij}\dot{x}^i\dot{x}^j$$

having suspended the summation convention, then the lemma states that for any $z' \in W \subset T(M)$, there exists a coordinate system on some neighborhood $W_{z'}$ of z' and a function H

such that

$$L\,(\,e_\sigma\,,A_p^\sigma\,,g_{pq}\,,\dot{x}^p\,)\;=\;H(\,e_\sigma\,,C_j^\alpha\,,e_{ij}\,)$$

Begin by considering an arbitrary point z' (coordinates $x'^h, \dot{x}'^h$) in W, and constructing a coordinate system in a neighborhood $W_{z'}$ of z' such that at the point z'

$$\boxed{\dot{x}'^h = 1} \tag{2.7}$$

(for $h = 1,...,n$). Such a neighborhood and coordinate system exist by virtue of the following construction. Consider the point $x' \in M$ which is the natural projection $x' = p(z')$ in M of the arbitrary point z', and any coordinate neighborhood U of x'. Since (1.8) requires that for some k, $\dot{x}'^k \neq 0$, we may by a rotation of coordinates on U induce a coordinate system on a neighborhood of z' such that $\dot{x}'^h \neq 0$ for $h = 1,...,n$. Next, we may scale each coordinate on U by an appropriate constant to induce a coordinate system in a neighborhood of z' such that (2.7) holds at z'.

For example, suppose at z' we have that

$$[\;\dot{x}'^1 \quad \dot{x}'^2 \quad \dot{x}'^3\;]^T = [\;3 \;\; 0 \;\; -1\,]^T$$

Then by a rotation in the $x'^1x'^2$-plane of $\pi/4$ radians, we obtain a coordinate system such that at z'

$$[\;\dot{\bar{x}}'^1 \quad \dot{\bar{x}}'^2 \quad \dot{\bar{x}}'^3\;]^T = [\;3/\sqrt{2} \;\; 3/\sqrt{2} \;\; -1\;]^T$$

A second change of coordinates, scaling the previous coordinates by $\sqrt{2}/3$, $\sqrt{2}/3$, and -1 respectively, determines a new coordinate system such that at z'

$$[\;\dot{\bar{x}}'^1 \quad \dot{\bar{x}}'^2 \quad \dot{\bar{x}}'^3\;]^T = [1 \quad 1 \quad 1]^T$$

Thus, we obtain a neighborhood of z', and a coordinate system on that neighborhood, such that (2.7) is satisfied at z'. To avoid unnecessary symbols, we will henceforth use unbarred coordinates to represent geometric objects in this coordinate system.

Now consider a neighborhood of an arbitrary point z' in W, and using the coordinate system just described recall the scalar function L of (1.7) -

$$L = L(e_\alpha, A_j^\alpha, g_{ij}, \dot{x}^j)$$

Consider also the same function, evaluated with related arguments

$$\boxed{\tilde{A}_j^\alpha = A_j^\alpha t^j \quad ; \quad \tilde{g}_{ij} = g_{ij} t^i t^j \quad ; \quad \dot{\tilde{x}}^j = \frac{\dot{x}^j}{t^j}} \tag{2.8}$$

replacing A_j^α, g_{ij}, $\dot{x}^j$ (in which we remember that the summation convention is not applied). These arguments in matrix form are

$$\begin{bmatrix} A_1^\alpha t^1 \\ A_2^\alpha t^2 \\ A_3^\alpha t^3 \\ \vdots \\ A_n^\alpha t^n \end{bmatrix} ; \quad \begin{bmatrix} g_{11}t^1t^1 & g_{12}t^1t^2 & g_{13}t^1t^3 & \cdots & g_{1n}t^1t^n \\ g_{21}t^2t^1 & g_{22}t^2t^2 & g_{23}t^2t^3 & \cdots & g_{2n}t^2t^n \\ g_{31}t^3t^1 & g_{32}t^3t^2 & g_{33}t^3t^3 & \cdots & g_{3n}t^3t^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{n1}t^nt^1 & g_{n2}t^nt^2 & g_{n3}t^nt^3 & \cdots & g_{nn}t^nt^n \end{bmatrix} ; \quad \begin{bmatrix} \dot{x}^1/t^1 \\ \dot{x}^2/t^2 \\ \dot{x}^3/t^3 \\ \vdots \\ \dot{x}^n/t^n \end{bmatrix}$$

and are equal to the usual arguments A_j^α, g_{ij}, $\dot{x}^j$ when $t^i = 1$ ($i = 1,\ldots,n$). The resulting function is indicated by the notation

$$\boxed{\widetilde{L} = L(e_\alpha, \widetilde{A}_j^\alpha, \widetilde{g}_{ij}, \dot{\widetilde{x}}^j)}$$

We have assumed an admissible domain (see section (1.2)) for L, so there exists for each $z' \subset W$ a positive real number $\varepsilon_{z'}$ and a neighborhood $W_{z'} \subset W$ of z, such that $\widetilde{L}$ is well defined for $t^i \in (1-\varepsilon_{z'}, 1+\varepsilon_{z'})$. We therefore introduce real-valued continuous functions t^i on $W_{z'}$:

$$\boxed{t^i = t^i(\dot{x}^h)}$$

whose range is some subset of $(1-\varepsilon_{z'}, 1+\varepsilon_{z'})$, and stipulate that $t^i = 1$ at z' (for $i = 1,\ldots,n$). The functions t^i are otherwise arbitrary.

If we substitute $\widetilde{L}$ into (2.3) and set $p = m$, we observe that for any point $z \in W_{z'}$ having coordinates x^i, $\dot{x}^i$

$$\sum_\alpha \frac{\partial \widetilde{L}}{\partial \widetilde{A}_m^\alpha} \widetilde{A}_m^\alpha - \frac{\partial \widetilde{L}}{\partial \dot{\widetilde{x}}^m} \dot{\widetilde{x}}^m + 2 \sum_k \frac{\partial \widetilde{L}}{\partial \widetilde{g}_{mk}} \widetilde{g}_{mk} = 0 \qquad (2.9)$$

Substitution from (2.8) then gives

$$\sum_\alpha \frac{\partial \widetilde{L}}{\partial \widetilde{A}_m^\alpha} t^m A_m^\alpha - \frac{\partial \widetilde{L}}{\partial \dot{\widetilde{x}}^m} \frac{\dot{x}^m}{t^m} + 2t^m \sum_k \frac{\partial \widetilde{L}}{\partial \widetilde{g}_{mk}} g_{mk} t^k = 0 \qquad (2.10)$$

Next define the function

$$h(t^j) = L - \widetilde{L} = L(e_\alpha, A_j^\alpha, g_{ij}, \dot{x}^j) - L(e_\alpha, \widetilde{A}_j^\alpha, \widetilde{g}_{ij}, \dot{\widetilde{x}}^j) \qquad (2.11)$$

(this function depends upon the usual arguments of L in addition to the functions t^i, but for convenience only the latter arguments are shown).

This function must be identically zero for the following reason. On the right side of (2.11), only the arguments $\widetilde{A}_j^\alpha$, $\widetilde{g}_{ij}$, $\dot{\widetilde{x}}^j$ of $\widetilde{L}$ depend on t^i. Therefore, by differentiation of both sides of (2.11) with respect to t^m we get

$$\frac{\partial h}{\partial t^m} = -\left(\sum_\alpha \frac{\partial \widetilde{L}}{\partial \widetilde{A}_m^\alpha} \frac{\partial \widetilde{A}_m^\alpha}{\partial t^m} + \sum_{i,j} \frac{\partial \widetilde{L}}{\partial \widetilde{g}_{ij}} \frac{\partial \widetilde{g}_{ij}}{\partial t^m} + \frac{\partial \widetilde{L}}{\partial \dot{\widetilde{x}}^m} \frac{\partial \dot{\widetilde{x}}^m}{\partial t^m} \right)$$

or

$$\frac{\partial h}{\partial t^m} = -\sum_\alpha \frac{\partial \widetilde{L}}{\partial \widetilde{A}_m^\alpha} A_m^\alpha - 2\sum_k \frac{\partial \widetilde{L}}{\partial \widetilde{g}_{mk}} g_{mk} t^k + \frac{\partial \widetilde{L}}{\partial \dot{\widetilde{x}}^m} \frac{\dot{x}^m}{(t^m)^2} \tag{2.12}$$

If (2.10) is divided by t^m and then substituted into (2.12), the latter becomes

$$\frac{\partial h}{\partial t^m} = 0$$

which implies that

$$h(t^j) = \text{constant}$$

But substitution of $t^i = 1$ into (2.11) shows that

$$h(1) = 0$$

so the constant must be zero, i.e.

$$h(t^j) = 0$$

Putting this result into (2.11) determines that at each point z in $W_{z'}$, and for any curve on M which is such that at $x = p(z)$ the tangent vector has components $\dot{x}^i$,

$$\boxed{L(e_\alpha, A_j^\alpha, g_{ij}, \dot{x}^j) = L(e_\alpha, \widetilde{A}_j^\alpha, \widetilde{g}_{ij}, \dot{\widetilde{x}}^j)} \tag{2.13}$$

Equation (2.13) holds for arbitrary continuous functions t^i whose ranges are each a subset of $(1-\varepsilon_{z'}, 1+\varepsilon_{z'})$, which are defined on $W_{z'}$, and which are such that $t^i = 1$ at z'.

As a result, if we let $t^m = \dot{x}^m$ in (2.13) and define

$$\boxed{C_j^\alpha = A_j^\alpha \dot{x}^j \ , \ e_{ij} = g_{ij}\,\dot{x}^i \dot{x}^j} \tag{2.14}$$

then (2.13) becomes

$$L(\, e_\sigma \, , C_j^\alpha \, , e_{ij} \, , 1\,) = L\,(\, e_\sigma \, , A_p^\sigma \, , g_{pq} \, , \dot{x}^p\,) \tag{2.15}$$

where "1" denotes replacement of the n arguments $\dot{x}^j$ by the number one. This establishes the following lemma.

Lemma 1: *Let L be a C^3 real-valued function of the quantities e_α, A_j^α, g_{ij}, $\dot{x}^j$ which is defined on an admissible domain. Let z' be an arbitrary point of the region W of the tangent bundle for which L is defined, and let C_j^α, e_{ij} denote the quantities (2.14). Then there exists a coordinate system on some neighborhood $W_{z'}$ of z', in terms of which a necessary condition in order that L be a scalar whose associated vector E_k is independent of $u_{j}{}^{\alpha}_{k}$, is that there exist a C^3 function H such that*

$$\boxed{L = H(\, e_\sigma \, , C_j^\alpha \, , e_{ij}\,)} \tag{2.16}$$

2.5 A Simple Functional Equation for the Lagrangian

We now show that in order that the vector E_k of (1.15) be independent of $u_{j}{}^{\alpha}_{k}$ (as required by (1.20)) it is necessary that there exist a C^3 scalar function L^* such that, whenever L is defined,

$$L = L^*(e_\alpha , \sum_{j=1}^{n} A_j^\alpha \dot{x}^j , \sum_{i,j=1}^{n} g_{ij} \dot{x}^i \dot{x}^j)$$

This result is derived as a consequence of Lemma 1, but does not require a special coordinate system. We continue to suspend use of the summation convention in this section.

Consider an arbitrary point $z' \in W$, and an associated neighborhood $W_{z'}$ and coordinate system as indicated in Lemma 1 (recall that (2.7) holds at the point z' in this coordinate system). The derivatives of the function H are related to the derivatives (1.21) of L by:

$$L_\alpha^j = \frac{\partial L}{\partial A_j^\alpha} = \frac{\partial H}{\partial C_j^\alpha} \dot{x}^j \tag{2.17}$$

$$L_k = \frac{\partial L}{\partial \dot{x}^k} = 2 \sum_{i=1}^{n} \frac{\partial H}{\partial e_{ik}} g_{ik} \dot{x}^i + \sum_{\beta=1}^{r} \frac{\partial H}{\partial C_k^\beta} A_k^\beta$$

$$\Lambda_k^\alpha = \frac{\partial L_k}{\partial e_\alpha} = 2 \sum_{i=1}^{n} \frac{\partial^2 H}{\partial e_\alpha \partial e_{ik}} g_{ik} \dot{x}^i + \sum_{\beta=1}^{r} \frac{\partial^2 H}{\partial e_\alpha \partial C_k^\beta} A_k^\beta$$

$$L_{\alpha k}^j = \frac{\partial L_k}{\partial A_j^\alpha} = 2 \sum_{i=1}^{n} \frac{\partial^2 H}{\partial C_j^\alpha \partial e_{ik}} g_{ik} \dot{x}^i \dot{x}^j + \sum_{\beta=1}^{r} \frac{\partial^2 H}{\partial C_j^\alpha \partial C_k^\beta} A_k^\beta \dot{x}^j + \frac{\partial H}{\partial C_k^\alpha} \delta_k^j \tag{2.18}$$

Equations (2.6) allow us to simplify the last expression. In terms of H, (2.6) is

$$\left(2\sum_{i=1}^{n}\frac{\partial^2 H}{\partial C_j^\alpha \partial e_{ik}} g_{ik}\dot{x}^i\dot{x}^j + \sum_{\beta=1}^{r}\frac{\partial^2 H}{\partial C_j^\alpha \partial C_k^\beta} A_k^\beta \dot{x}^j + \frac{\partial H}{\partial C_k^\alpha}\delta_k^j\right)\dot{x}^j - \left(\frac{\partial H}{\partial C_j^\alpha}\dot{x}^j\right)\delta_k^j = 0$$

or, after cancellations

$$2\sum_{i=1}^{n}\frac{\partial^2 H}{\partial C_j^\alpha \partial e_{ik}} g_{ik}\dot{x}^i\dot{x}^j + \sum_{\beta=1}^{r}\frac{\partial^2 H}{\partial C_j^\alpha \partial C_k^\beta} A_k^\beta \dot{x}^j = 0 \qquad (2.19)$$

By substitution of (2.19) into the expression (2.18) for $L_{\alpha k}^j$, we deduce that

$$L_{\alpha k}^j = \frac{\partial H}{\partial C_k^\alpha}\delta_k^j \qquad (2.20)$$

These expressions for the derivatives of H allow us to express (2.5) in terms of H. Substitute (2.17) and (2.20) into (2.5), to obtain

$$\left[\left(\frac{\partial H}{\partial C_k^\alpha}\delta_k^j\right)\dot{x}^m - \left(\frac{\partial H}{\partial C_j^\alpha}\dot{x}^j\right)\delta_k^m\right] + \left[\left(\frac{\partial H}{\partial C_k^\alpha}\delta_k^m\right)\dot{x}^j - \left(\frac{\partial H}{\partial C_m^\alpha}\dot{x}^m\right)\delta_k^j\right] = 0$$

Then, in the special case $k = m'$ and $j = m'$ (where m is any (fixed) index and m' represents an arbitrary index not equal to m) we have

$$\left[\frac{\partial H}{\partial C_{m'}^\alpha}\dot{x}^m\right] + \left[-\frac{\partial H}{\partial C_m^\alpha}\dot{x}^m\right] = 0$$

This system of equations may be solved as follows. Since we are in a coordinate system for which $\dot{x}^m \neq 0$, therefore

$$\frac{\partial H}{\partial C^{\alpha}_{m'}} - \frac{\partial H}{\partial C^{\alpha}_{m}} = 0 \tag{2.21}$$

In (2.21), replace m' by m'' (another arbitrary index not equal to m) to get

$$\frac{\partial H}{\partial C^{\alpha}_{m''}} - \frac{\partial H}{\partial C^{\alpha}_{m}} = 0$$

and subtract this result from (2.21) to obtain (for any $m' \neq m$, and $m'' \neq m$)

$$\frac{\partial H}{\partial C^{\alpha}_{m'}} - \frac{\partial H}{\partial C^{\alpha}_{m''}} = 0 \tag{2.22}$$

Equations (2.21) and (2.22) together imply that for any indices α, h, k

$$\frac{\partial H}{\partial C^{\alpha}_{h}} - \frac{\partial H}{\partial C^{\alpha}_{k}} = 0 \tag{2.23}$$

Accordingly, equation (A.2) from the appendix indicates that H can only depend on the quantities C^{α}_{k} through their sum. That is, if we define

$$\boxed{A^{\alpha} = \sum_{j=1}^{n} C^{\alpha}_{j} = \sum_{j=1}^{n} A^{\alpha}_{j} \dot{x}^{j}} \tag{2.24}$$

then, for the neighborhood $W_{z'}$ of z', there must exist a function $\widetilde{H}$ such that

$$\boxed{H = \widetilde{H}(e_{\alpha}, A^{\alpha}, e_{ij})} \tag{2.25}$$

This result can be simplified further, by use of (2.3). Since by (2.16) and (2.25)

$$\widetilde{H} = H = L$$

therefore $\widetilde{H}$ must satisfy (2.3):

$$\sum_{\alpha=1}^{r} \frac{\partial L}{\partial A_m^\alpha} A_p^\alpha - \frac{\partial L}{\partial \dot{x}^p} \dot{x}^m + 2 \sum_{k=1}^{n} \frac{\partial L}{\partial g_{mk}} g_{pk} = 0$$

that is

$$\left(\sum_\alpha \frac{\partial \widetilde{H}}{\partial A^\alpha}\right) A_p^\alpha \dot{x}^m - \left(\sum_\alpha \frac{\partial \widetilde{H}}{\partial A^\alpha} A_p^\alpha + 2\sum_k \frac{\partial \widetilde{H}}{\partial e_{pk}} g_{pk} \dot{x}^k\right) \dot{x}^m + 2\left(\sum_k \frac{\partial \widetilde{H}}{\partial e_{mk}} \dot{x}^m \dot{x}^k\right) g_{pk} = 0$$

(where e_{pk} is defined by (2.14)). After some cancellations we have

$$-\left(2\sum_k \frac{\partial \widetilde{H}}{\partial e_{pk}} g_{pk} \dot{x}^k\right) \dot{x}^m + \left(2\sum_k \frac{\partial \widetilde{H}}{\partial e_{mk}} \dot{x}^m \dot{x}^k\right) g_{pk} = 0$$

or

$$\sum_k \left(\frac{\partial \widetilde{H}}{\partial e_{mk}} - \frac{\partial \widetilde{H}}{\partial e_{pk}}\right) g_{pk} \dot{x}^m \dot{x}^k = 0 \tag{2.26}$$

If m is replaced by j in this equation we find that

$$\sum_k \left(\frac{\partial \widetilde{H}}{\partial e_{jk}} - \frac{\partial \widetilde{H}}{\partial e_{pk}}\right) g_{pk} \dot{x}^j \dot{x}^k = 0 \tag{2.27}$$

Next we multiply (2.26) by $\dot{x}^j$, multiply (2.27) by $\dot{x}^m$, and subtract. The result is

$$\sum_k \left(\frac{\partial \widetilde{H}}{\partial e_{mk}} - \frac{\partial \widetilde{H}}{\partial e_{jk}}\right) g_{pk} \dot{x}^m \dot{x}^j \dot{x}^k = 0 \tag{2.28}$$

Since in our special coordinate neighborhood $\dot{x}^i \neq 0$ (for i = 1,...,n), we may multiply

(2.28) by $\frac{\dot{x}^p}{\dot{x}^m \dot{x}^j}$ to obtain

$$\sum_k \left(\frac{\partial \widetilde{H}}{\partial e_{mk}} - \frac{\partial \widetilde{H}}{\partial e_{jk}} \right) e_{pk} = 0 \tag{2.29}$$

But according to (2.14), $[e_{pk}]$ is the matrix

$$\begin{bmatrix} \dot{x}^1 & 0 & \cdots & 0 \\ 0 & \dot{x}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \dot{x}^n \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & & g_{2n} \\ \vdots & & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix} \begin{bmatrix} \dot{x}^1 & 0 & \cdots & 0 \\ 0 & \dot{x}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \dot{x}^n \end{bmatrix}$$

and therefore,

$$\det(e_{ij}) = (\dot{x}^1 \dot{x}^2 \ldots \dot{x}^n)^2 \det(g_{ij}) \neq 0$$

Hence, (2.29) is equivalent to

$$\frac{\partial \widetilde{H}}{\partial e_{mk}} - \frac{\partial \widetilde{H}}{\partial e_{pk}} = 0$$

According to equation (A.2) (in the appendix) the validity of this system of equations requires that $\widetilde{H}$ is such that for some function L^* (defined for all z in $W_{z'}$ and all corresponding curves on M)

$$\widetilde{H} = L^*(e_\alpha , A^\alpha , \gamma)$$

in which the A^α are given by (2.24) and where we have written

$$\boxed{\gamma = \sum_{i,j=1}^{n} e_{ij} = \sum_{i,j=1}^{n} g_{ij}\,\dot{x}^i\,\dot{x}^j} \tag{2.30}$$

Because of (2.24) and (2.30), the arguments of L^* are scalars, so *this result holds not only in our special coordinate system but in arbitrary coordinate systems* on $W_{z'}$.

Since by (2.16) and (2.25) we know that $\widetilde{H} = L$, it has been shown that corresponding to any point z' in the admissible region of the tangent bundle, there exists a scalar function L^* such that for a neighborhood $W_{z'}$ of z' L admits the representation

$$L = L^*(e_\alpha, A^\alpha, \gamma) \tag{2.31}$$

(in which the arguments A^α and γ are defined by (2.24) and (2.30)). Since L is C^3, it follows that L^* is also C^3. This establishes the following theorem (for which the summation convention is resumed).

Theorem 1: *Let L be a C^3 function of the quantities e_α, A_j^α, g_{ij}, $\dot{x}^j$ which is defined on an admissible domain. Also, let*

$$A^\alpha = A_j^\alpha\,\dot{x}^j \quad ; \quad \gamma = g_{ij}\,\dot{x}^i\,\dot{x}^j$$

Then in order that L be a scalar which does not depend upon the parameters $u_h{}^\alpha{}_k$, it is necessary that there exist a C^3 scalar function L^* such that

$$\boxed{L = L^*(e_\alpha, A^\alpha, \gamma)} \tag{2.32}$$

2.6 The Parameter-Invariant Variational Principle

When the assumption that the fundamental integral (1.11) is parameter-invariant is applied to this result an even stronger result is obtained (which constitutes a corollary to Theorem 1). We maintain the summation convention.

The parameter-invariance of the fundamental integral was found to be equivalent to the conditions (1.12) and (1.13). Let us express these equations in terms of the function L^* of Theorem 1. Equation (1.12) becomes

$$\left[\frac{\partial \mathring{L}^*}{\partial \mathring{A}^\alpha} \mathring{A}^\alpha_k + 2\frac{\partial \mathring{L}^*}{\partial \gamma} g_{ik}\dot{x}^i\right]\dot{x}^k = \mathring{L}^*$$

that is,

$$\boxed{\frac{\partial \mathring{L}^*}{\partial \mathring{A}^\alpha} \mathring{A}^\alpha + 2\frac{\partial \mathring{L}^*}{\partial \gamma}\gamma = \mathring{L}^*} \tag{2.33}$$

Since (2.32) requires that $\mathring{L} = \mathring{L}^*$, we may substitute the left-hand side of (2.33) for the function $\mathring{L}$ in (1.13), obtaining

$$\boxed{\left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \mathring{A}^\alpha}\mathring{A}^\alpha + 2\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \gamma}\gamma\right]\mathring{f}_\sigma = 0} \tag{2.34}$$

Thus, under the assumptions of Theorem 1, *equations (2.33) and (2.34) constitute a set of necessary and sufficient conditions for the fundamental integral to be parameter-invariant.*

A second approach to the equation (1.12) also yields a useful result. Since we are restricted to a region of the tangent bundle for which $\gamma \neq 0$, the relations

$$\xi = \sqrt{|\gamma|} \quad ; \quad B^{\alpha} = \frac{A^{\alpha}}{\sqrt{|\gamma|}}$$

represent an invertible change of variables. We deduce therefore via (2.32), that there exists a function L^{**} such that

$$L = L^{**}(e_{\alpha}, B^{\alpha}, \xi)$$

Since ξ and B^{α} are homogeneous of degree one and zero in $\dot{x}^i$ respectively, it follows from (1.12) that

$$\frac{\partial \mathring{L}^{**}}{\partial \xi} \xi = \mathring{L}^{**}$$

which implies that there exists a scalar function $\Lambda(\mathring{e}_{\alpha}, \mathring{B}^{\alpha})$, such that

$$\mathring{L} = \xi \Lambda(\mathring{e}_{\alpha}, \mathring{B}^{\alpha}) \tag{2.35}$$

Hence have shown that under the conditions of Theorem 1, equation (1.12) implies (2.35). Differentiation of (2.35) however, shows that it in turn implies (1.12). Therefore we have established the following corollary to Theorem 1.

Corollary 1: *If the conditions of Theorem 1 are satisfied, then the condition*

$$\frac{\partial \mathring{L}}{\partial \dot{x}^k} \dot{x}^k = \mathring{L}$$

is valid if and only if there exists a C^3 scalar function $\Lambda(\mathring{e}_{\alpha}, \frac{\mathring{A}^{\alpha}}{\sqrt{|\gamma|}})$ such that

$$\boxed{\mathring{L} = \sqrt{|\gamma|}\, \Lambda(\mathring{e}_{\alpha}, \frac{\mathring{A}^{\alpha}}{\sqrt{|\gamma|}})} \tag{2.36}$$

CHAPTER 3

THE EULER-LAGRANGE EQUATIONS REVISITED

We now endeavor to find an expression for the Euler-Lagrange equations, which takes into account the results of chapter 2. For convenience, we denote by $\overset{\circ}{L}{}^*$ the function

$$\overset{\circ}{L}{}^* = L^*(e_\alpha(\tau, u^\sigma(x^h)), A_j^\alpha(x^h, u^\sigma(x^h), u_j^\sigma(x^h)), g_{ij}(x^h)\dot{x}^i\dot{x}^j)$$

where L^* is the function (2.32).

3.1 A Required Form for the Euler-Lagrange Vector

The derivatives required in the Euler-Lagrange vector (1.14) are:

$$\frac{\partial \overset{\circ}{L}{}^*}{\partial x^k} = \frac{\partial \overset{\circ}{L}{}^*}{\partial \overset{\circ}{A}{}^\alpha}\frac{\partial \overset{\circ}{A}{}^\alpha}{\partial x^k} + \frac{\partial \overset{\circ}{L}{}^*}{\partial \gamma}\frac{\partial \gamma}{\partial x^k}$$

$$= \frac{\partial \overset{\circ}{L}{}^*}{\partial \overset{\circ}{A}{}^\alpha}\overset{\circ}{A}{}^\alpha_{j,k}\dot{x}^j + \frac{\partial \overset{\circ}{L}{}^*}{\partial \gamma}g_{ij,k}\dot{x}^i\dot{x}^j \tag{3.1}$$

and

$$\frac{\partial \overset{\circ}{L}{}^*}{\partial \dot{x}^k} = \frac{\partial \overset{\circ}{L}{}^*}{\partial \overset{\circ}{A}{}^\alpha}\overset{\circ}{A}{}^\alpha_k + \frac{\partial \overset{\circ}{L}{}^*}{\partial \gamma}\frac{\partial \gamma}{\partial \dot{x}^k}$$

$$= \frac{\partial \overset{\circ}{L}{}^*}{\partial \overset{\circ}{A}{}^\alpha}\overset{\circ}{A}{}^\alpha_k + 2\frac{\partial \overset{\circ}{L}{}^*}{\partial \gamma}g_{ik}\dot{x}^i \tag{3.2}$$

and

$$\frac{d}{d\tau}\left(\frac{\partial \mathring{L}^*}{\partial \dot{x}^k}\right) = \frac{d}{d\tau}\left(\frac{\partial \mathring{L}^*}{\partial \mathring{A}^\alpha}\mathring{A}_k^\alpha\right) + 2\frac{d}{d\tau}\left(\frac{\partial \mathring{L}^*}{\partial \gamma}g_{ik}\dot{x}^i\right)$$

$$= \left(\left[\frac{\partial^2 \mathring{L}^*}{\partial e_\sigma \partial \mathring{A}^\alpha}\mathring{f}_\sigma + \frac{\partial^2 \mathring{L}^*}{\partial \mathring{A}^\sigma \partial \mathring{A}^\alpha}\left(\mathring{A}_j^\sigma \ddot{x}^j + \mathring{A}_{j,p}^\sigma \dot{x}^j \dot{x}^p\right) + \frac{\partial^2 \mathring{L}^*}{\partial \gamma \partial \mathring{A}^\alpha}\frac{d\gamma}{d\tau}\right]\mathring{A}_k^\alpha + \frac{\partial \mathring{L}^*}{\partial \mathring{A}^\alpha}\mathring{A}_{k,j}^\alpha \dot{x}^j\right)$$

$$+ 2\left(\left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \gamma}f_\sigma + \frac{\partial^2 \mathring{L}^*}{\partial \mathring{A}^\sigma \partial \gamma}\left(\mathring{A}_j^\sigma \ddot{x}^j + \mathring{A}_{j,p}^\sigma \dot{x}^j \dot{x}^p\right) + \frac{\partial^2 \mathring{L}^*}{\partial \gamma \partial \gamma}\frac{d\gamma}{d\tau}\right]g_{ik}\dot{x}^i + \frac{\partial \mathring{L}^*}{\partial \gamma}\left(g_{ik,j}\dot{x}^j \dot{x}^i + g_{ik}\ddot{x}^i\right)\right)$$

(3.3)

Some of the terms in (3.3) cancel. To see this, notice that (2.6) indicates that (briefly suspending the summation convention)

$$\frac{\partial^2 L}{\partial A_j^\alpha \partial \dot{x}^k}\dot{x}^j - \frac{\partial L}{\partial A_j^\alpha}\delta_k^j = 0$$

which is, in terms of L^*

$$\left[\sum_{i=1}^{n}\dot{x}^j\left(\frac{\partial^2 L^*}{\partial A^\alpha \partial \gamma}\right)2\, g_{ik}\dot{x}^i + \sum_{\beta=1}^{r}\dot{x}^j\left(\frac{\partial^2 L^*}{\partial A^\alpha \partial A^\beta}\right)A_k^\beta + \frac{\partial L^*}{\partial A^\alpha}\delta_k^j\right]\dot{x}^j - \frac{\partial L^*}{\partial A^\alpha}\dot{x}^j\,\delta_k^j = 0$$

Two terms cancel, and we divide the result by $\dot{x}^j$ (since (1.8) shows that there exists a j for which $\dot{x}^j \neq 0$) to get

$$2\sum_{i=1}^{n}\left(\frac{\partial^2 L^*}{\partial A^\alpha \partial \gamma}\right)g_{ik}\dot{x}^i + \sum_{\beta=1}^{r}\left(\frac{\partial^2 L^*}{\partial A^\alpha \partial A^\beta}\right)A_k^\beta = 0$$

We restore the summation convention and have

$$\boxed{2\frac{\partial^2 L^*}{\partial A^{\alpha}\partial\gamma}g_{ik}\dot{x}^i + \frac{\partial^2 L^*}{\partial A^{\alpha}\partial A^{\beta}}A_k^{\beta} = 0} \tag{3.4}$$

Substitution of (3.4) (after evaluation on the identity section) into (3.3) reduces the latter to

$$\frac{d}{d\tau}\left(\frac{\partial \mathring{L}^*}{\partial \dot{x}^k}\right) = \left(\left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_{\sigma}\partial \mathring{A}^{\alpha}}\mathring{f}_{\sigma} + \frac{\partial^2 \mathring{L}^*}{\partial g\,\partial \mathring{A}^{\alpha}}\frac{d\gamma}{d\tau}\right]\mathring{A}_k^{\alpha} + \frac{\partial \mathring{L}^*}{\partial \mathring{A}^{\alpha}}\mathring{A}^{\alpha}_{k,j}\dot{x}^j\right)$$
$$+ 2\left(\left[\frac{\partial^2 \mathring{L}^*}{\partial e_{\sigma}\partial\gamma}f_{\sigma} + \frac{\partial^2 \mathring{L}^*}{\partial\gamma\partial\gamma}\frac{d\gamma}{d\tau}\right]g_{ik}\dot{x}^i + \frac{\partial \mathring{L}^*}{\partial\gamma}\left(g_{ik,j}\dot{x}^j\dot{x}^i + g_{ik}\ddot{x}^i\right)\right) \tag{3.5}$$

Thus, by means of (3.1) and (3.5) the Euler-Lagrange vector may be expressed as

$$\mathring{E}_k = \frac{d}{d\tau}\left(\frac{\partial \mathring{L}^*}{\partial \dot{x}^k}\right) - \frac{\partial \mathring{L}^*}{\partial x^k}$$

or

$$\mathring{E}_k = \left(\left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_{\sigma}\partial \mathring{A}^{\alpha}}\mathring{f}_{\sigma} + \frac{\partial^2 \mathring{L}^*}{\partial\gamma\partial \mathring{A}^{\alpha}}\frac{d\gamma}{d\tau}\right]\mathring{A}_k^{\alpha} + \frac{\partial \mathring{L}^*}{\partial \mathring{A}^{\alpha}}\mathring{A}^{\alpha}_{k,j}\dot{x}^j\right)$$
$$+ 2\left(\left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_{\sigma}\partial\gamma}\mathring{f}_{\sigma} + \frac{\partial^2 \mathring{L}^*}{\partial\gamma\partial\gamma}\frac{d\gamma}{d\tau}\right]g_{ik}\dot{x}^i + \frac{\partial \mathring{L}^*}{\partial\gamma}\left(g_{ik,j}\dot{x}^j\dot{x}^i + g_{ik}\ddot{x}^i\right)\right)$$
$$- \left(\frac{\partial \mathring{L}^*}{\partial \mathring{A}^{\alpha}}\mathring{A}^{\alpha}_{j,k}\dot{x}^j + \frac{\partial \mathring{L}^*}{\partial\gamma}g_{ij,k}\dot{x}^i\dot{x}^j\right)$$

which is, due to (0.18) and (0.19)

$$\mathring{E}_k = 2\frac{\partial \mathring{L}^*}{\partial \gamma} g_{ik}\frac{D\dot{x}^i}{D\tau} + \left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \mathring{A}^\alpha}\mathring{A}^\alpha_k + 2\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \gamma} g_{ik}\dot{x}^i\right]\mathring{f}_\sigma + \frac{\partial \mathring{L}^*}{\partial \mathring{A}^\alpha}\left(\mathring{A}^\alpha_{k,j} - \mathring{A}^\alpha_{j,k}\right)\dot{x}^j + \left[\frac{\partial^2 \mathring{L}^*}{\partial \gamma \partial \mathring{A}^\alpha}\mathring{A}^\alpha_k + 2\frac{\partial^2 \mathring{L}^*}{\partial \gamma \partial \gamma} g_{ik}\dot{x}^i\right]\frac{d\gamma}{d\tau}$$

If the quantities in the Euler-Lagrange expression which (as indicated by the superscript "o") have been evaluated at $u = e$ are instead evaluated for arbitrary u, then the vector E_k of (1.15) is obtained - namely

$$E_k = 2\frac{\partial L^*}{\partial \gamma} g_{ik}\frac{D\dot{x}^i}{D\tau} + \left[\frac{\partial^2 L^*}{\partial e_\sigma \partial A^\alpha}A^\alpha_k + 2\frac{\partial^2 L^*}{\partial e_\sigma \partial \gamma} g_{ik}\dot{x}^i\right]f_\sigma + \frac{\partial L^*}{\partial A^\alpha}\left(A^\alpha_{k,j} - A^\alpha_{j,k}\right)\dot{x}^j + \left[\frac{\partial^2 L^*}{\partial \gamma \partial A^\alpha}A^\alpha_k + 2\frac{\partial^2 L^*}{\partial \gamma \partial \gamma} g_{ik}\dot{x}^i\right]\frac{d\gamma}{d\tau} \tag{3.6}$$

3.2 The Existence and Uniqueness of Solutions of the Euler-Lagrange Equations

Because the Lagrangian is positively homogeneous of degree one (due to (1.12)) and does not depend explicitly on the parameter (due to (1.13)) the Euler-Lagrange equations are invariant under an admissible change of parameter.[14] We may therefore take τ to be a constant multiple of the proper time parameter (0.13). In this case γ is a constant and

[14] This is a standard result whose proof may be found (for example) in Rund [11], p. 172.

$$\frac{d\gamma}{d\tau} = 0$$

so that (in this special case) (3.6) is reduced to

$$\boxed{E_k = \left(2\frac{\partial L^*}{\partial \gamma}\right) g_{ik}\frac{D\dot{x}^i}{D\tau} + \left(\frac{\partial^2 L^*}{\partial e_\sigma \partial A^\alpha} A^\alpha_k + 2\frac{\partial^2 L^*}{\partial e_\sigma \partial \gamma} g_{ik}\dot{x}^i\right) f_\sigma - \frac{\partial L^*}{\partial A^\alpha}(A^\alpha_{j,k} - A^\alpha_{k,j})\dot{x}^j} \tag{3.7}$$

Since according to (1.9) (and the fact that for some i, $\dot{x}^i \neq 0$)

$$\frac{\partial L^*}{\partial \gamma} \neq 0 \tag{3.8}$$

we may define functions M^*_k by writing

$$2\frac{\partial L^*}{\partial \gamma} M^*_k = \frac{\partial L^*}{\partial A^\alpha}\left(A^\alpha_{j,k} - A^\alpha_{k,j}\right)\dot{x}^j - \left[\frac{\partial^2 L^*}{\partial e_\sigma \partial A^\alpha} A^\alpha_k + 2\frac{\partial^2 L^*}{\partial e_\sigma \partial \gamma} g_{ik}\dot{x}^i\right] f_\sigma \tag{3.9}$$

and express equations (3.7) as

$$E_k = 2\frac{\partial L^*}{\partial \gamma}\left(g_{ik}\frac{D\dot{x}^i}{D\tau} - M^*_k\right) \tag{3.10}$$

Therefore the Euler-Lagrange equations ($\mathring{E}_k = 0$) become

$$\frac{D\dot{x}^i}{D\tau} = g^{ki}\mathring{M}^*_k \tag{3.11}$$

which is a system of $2n$ first order ordinary differential equations:

$$\frac{dx^i}{d\tau} = \dot{x}^i \quad ; \quad \frac{d\dot{x}^i}{d\tau} = g^{ki}\mathring{M}^*_k - \Gamma^i_{jk}\dot{x}^j\dot{x}^k$$

According to a standard theorem[15] from the theory of differential equations, there exists a real number $\varepsilon > 0$, such that this system of equations, referred to given parameter τ, with prescribed initial values in the domain of L

$$x^h(\tau_o) = x_o^h \quad ; \quad \frac{dx^h}{d\tau}(\tau_o) = \dot{x}_o^h \tag{3.12}$$

has a unique solution for $\tau \in (\tau_0 - \varepsilon, \tau_0 + \varepsilon)$. Hence there exists an $\varepsilon > 0$ such that equations (3.7) together with the initial condition (3.12) have a unique solution for $\tau \in (\tau_0 - \varepsilon, \tau_0 + \varepsilon)$. We shall show that this solution is such that γ is a constant of the motion.

In order that γ remain constant along a given curve, it is necessary and sufficient that (0.20) hold:

$$g_{ij}\,\dot{x}^i \frac{D\dot{x}^j}{D\tau} = 0$$

For a solution of (3.11), this is

$$g_{ij}\,\dot{x}^i g^{kj} \mathring{M}_k^* = 0$$

i.e.

$$\mathring{M}_k^* \dot{x}^k = 0$$

15 The theorem applies to systems of ordinary differential equations

$$\frac{dy^I}{dt} = f^I(t, y^J) \qquad (I,J = 1,\dots,N)$$

for which each f^I is defined on an open subset of R^{N+1}, is continuous and has continuous first partials

$$\frac{\partial f^I}{\partial y^J}$$

on that region. Such a system, together with an initial condition $y^I(t_o) = y_o^I$ is known to have a unique solution for t in some interval about t_0. See Pontryagin [12], p. 18, for a statement of this theorem.

Equation (3.9) shows that this condition is equivalent to

$$\frac{\partial \mathring{L}^*}{\partial \mathring{A}^\alpha}\left(\mathring{A}^\alpha_{k,j} - \mathring{A}^\alpha_{j,k}\right)\dot{x}^j\dot{x}^k - \left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \mathring{A}^\alpha}\mathring{A}^\alpha + 2\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \gamma}\gamma\right]\mathring{f}_\sigma = 0$$

that is

$$\left[\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \mathring{A}^\alpha}\mathring{A}^\alpha + 2\frac{\partial^2 \mathring{L}^*}{\partial \mathring{e}_\sigma \partial \gamma}\gamma\right]\mathring{f}_\sigma = 0$$

But this is equation (2.34), which is known to be valid as a result of the parameter-invariance of the fundamental integral.

Hence, we have shown that when τ is taken to be a constant multiple of proper time the Euler-Lagrange equations have a unique (local) solution. But we have also noted that the Euler-Lagrange equations are invariant under admissible changes of parameter. Therefore the Euler-Lagrange equations have a unique parameter-independent solution.

Corollary 2: *In addition to the conditions of theorem 1, let the fundamental integral (1.11) be parameter-invariant, and let the inequality (1.9) be satisfied. For admissible initial values (3.12), the resulting Euler-Lagrange equations together with these initial values have a unique solution on an interval $\tau \in (\tau_o - \varepsilon, \tau_o + \varepsilon)$ for some real number $\varepsilon > 0$.*

CHAPTER 4

GAUGE-INVARIANCE REQUIREMENTS

The functions E_p in (3.6) depend upon the quantities $e_\alpha, f_\alpha, A^\alpha_j, A^\alpha_{j,h}, \dot{x}^j, \ddot{x}^j, g_{jh}, g_{jh,k}$. In this chapter we consider general C^1 functions of such arguments

$$\boxed{M_p = M_p\,(\,e_\alpha\,, f_\alpha\,, A^\alpha_j\,, A^\alpha_{j,h}\,, \dot{x}^j\,, \ddot{x}^j\,, g_{jh}\,, g_{jh,k}\,)} \tag{4.1}$$

and deduce necessary and sufficient conditions for their gauge-invariance.

4.1 The First Invariance Identity

Recall that the functions (4.1) are said to be gauge-invariant if

$$M_p\,(e_\alpha, f_\alpha, A^\alpha_j, A^\alpha_{j,h}\,, \dot{x}^j, \ddot{x}^j, g_{jh}\,, g_{jh,k}\,) \;=\; M_p\,(\mathring{e}_\alpha, \mathring{f}_\alpha, \mathring{A}^\alpha_j, \mathring{A}^\alpha_{j,h}\,, \dot{x}^j, \ddot{x}^j, g_{jh}\,, g_{jh,k}\,) \tag{4.2}$$

Therefore, if we define

$$M_p =$$

$$M_p(e_\alpha(\tau,u^\sigma), f_\alpha(\tau,\dot{x}^h,u^\sigma,u^\rho_i), A^\alpha_j(x^i,u^\sigma,u^\rho_i), A^\alpha_{j,h}(x^i,u^\sigma,u^\rho_i,u^\sigma_{ij}), \dot{x}^j, \ddot{x}^i, g_{jh}(x^i), g_{hj,k}(x^i))$$

then in analogy to (1.18), (1.19) and (1.20), we see that the functions M_p are gauge-invariant if and only if the three sets of equations

$$\boxed{\frac{\partial M_p}{\partial u_{k\,q}^{\varepsilon}} = 0 \quad ; \quad \frac{\partial M_p}{\partial u_k^{\varepsilon}} = 0 \quad ; \quad \frac{\partial M_p}{\partial u^{\varepsilon}} = 0} \qquad (4.3)$$

are satisfied. If we express these sets of equations in terms of partial derivatives of M_p

$$\boxed{M_p^{\alpha} = \frac{\partial M_p}{\partial e_{\alpha}} \quad ; \quad N_p^{\alpha} = \frac{\partial M_p}{\partial f_{\alpha}} \quad ; \quad M_{p\alpha}^{j} = \frac{\partial M_p}{\partial A_j^{\alpha}} \quad ; \quad M_{p\alpha}^{jh} = \frac{\partial M_p}{\partial A_{j,h}^{\alpha}}}$$

the resulting three sets of equations shall be called the first, second and third *gauge-invariance identities* respectively.

Since $A_{j,h}^{\alpha}$ is the only argument of M_p which depends on $u_{k\,q}^{\varepsilon}$, the first set of equations from (4.3) is

$$M_{p\alpha}^{jh} \frac{\partial A_{j,h}^{\alpha}}{\partial u_{k\,q}^{\varepsilon}} = 0$$

which, due to (0.50), is

$$- M_{p\alpha}^{jh} \lambda_{\beta}^{\alpha} \frac{\partial u_{j\,h}^{\beta}}{\partial u_{k\,q}^{\varepsilon}} = 0$$

or

$$- M_{p\alpha}^{jh} \frac{1}{2} \lambda_{\beta}^{\alpha} \delta_{\varepsilon}^{\beta} (\delta_j^k \delta_h^q + \delta_h^k \delta_j^q) = 0$$

By application of (0.30), we obtain the <u>first invariance identity</u>

$$\boxed{M_{p\varepsilon}^{kq} + M_{p\varepsilon}^{qk} = 0} \qquad (4.4)$$

4.2 The Second Invariance Identity

The second set of equations in (4.3) may be written as

$$N_p^\alpha \frac{\partial f_\alpha}{\partial u_k^\xi} + M_{p\alpha}^{j} \frac{\partial A_j^\alpha}{\partial u_k^\xi} + M_{p\alpha}^{jh} \frac{\partial A_{j,h}^\alpha}{\partial u_k^\xi} = 0$$

From (1.4), (0.46) and (0.50), the derivatives shown are

$$\frac{\partial f_\alpha}{\partial u_k^\xi} = -e_\mu C_{\gamma\alpha}^{\mu} \lambda_\varepsilon^\gamma \dot{x}^k$$

$$\frac{\partial A_j^\alpha}{\partial u_k^\xi} = -\lambda_\varepsilon^{\ \alpha} \delta_j^k$$

and

$$\frac{\partial A_{j,h}^\alpha}{\partial u_k^\xi} = \frac{\partial G_\beta^\alpha}{\partial u^\varepsilon} \mathring{A}_j^\beta \delta_h^{\ k} - \frac{\partial \lambda_\beta^\alpha}{\partial u^\varepsilon} u_j^\beta \delta_h^{\ k} - \frac{\partial \lambda_\varepsilon^\alpha}{\partial u^\gamma} u_h^\gamma \delta_j^{\ k}$$

so we have

$$-N_p^\alpha\left(e_\mu C_{\gamma\alpha}^{\mu} \lambda_\varepsilon^\gamma \dot{x}^k\right) - M_{p\alpha}^{k} \lambda_\varepsilon^{\alpha} + \left(\frac{\partial G_\beta^\alpha}{\partial u^\varepsilon} \mathring{A}_j^\beta M_{p\alpha}^{jk} - \frac{\partial \lambda_\beta^\alpha}{\partial u^\varepsilon} u_j^\beta M_{p\alpha}^{jk} - \frac{\partial \lambda_\varepsilon^\alpha}{\partial u^\gamma} u_h^\gamma M_{p\alpha}^{kh}\right) = 0$$

By use of (4.4) this may be rewritten as

$$N_p^\alpha\left(e_\mu C_{\gamma\alpha}^{\mu} \lambda_\varepsilon^\gamma \dot{x}^k\right) + M_{p\alpha}^{k} \lambda_\varepsilon^{\alpha} - \frac{\partial G_\beta^\alpha}{\partial u^\varepsilon} \mathring{A}_j^\beta M_{p\alpha}^{jk} + \left(\frac{\partial \lambda_\beta^\alpha}{\partial u^\varepsilon} - \frac{\partial \lambda_\varepsilon^\alpha}{\partial u^\beta}\right) u_j^\beta M_{p\alpha}^{jk} = 0$$

and by additional application of (0.35) and (0.36) we obtain

$$N_p^\alpha\left(e_\mu C_{\gamma\alpha}^{\mu}\lambda_\varepsilon^\gamma \dot{x}^k\right) + M_{p\alpha}^k \lambda_\varepsilon^\alpha - C_{\eta\mu}^{\alpha}\lambda_\varepsilon^\eta G_\beta^\mu \mathring{A}_j^\beta M_{p\alpha}^{jk} + C_{\eta\mu}^{\alpha}\lambda_\varepsilon^\eta \lambda_\beta^\mu u_j^\beta M_{p\alpha}^{jk} = 0$$

or

$$N_p^\alpha\left(e_\mu C_{\gamma\alpha}^{\mu}\lambda_\varepsilon^\gamma \dot{x}^k\right) + M_{p\alpha}^k \lambda_\varepsilon^\alpha - C_{\eta\mu}^{\alpha}\lambda_\varepsilon^\eta G_\beta^\mu \mathring{A}_j^\beta M_{p\alpha}^{jk} + C_{\eta\mu}^{\alpha}\lambda_\varepsilon^\eta \left(G_\gamma^\mu \mathring{A}_j^\gamma - A_j^\mu\right) M_{p\alpha}^{jk} = 0$$

After some evident cancellations, there results

$$N_p^\alpha\left(e_\mu C_{\gamma\alpha}^{\mu}\lambda_\varepsilon^\gamma \dot{x}^k\right) + M_{p\alpha}^k \lambda_\varepsilon^\alpha + M_{p\alpha}^{kj} C_{\eta\mu}^{\alpha}\lambda_\varepsilon^\eta A_j^\mu = 0$$

or

$$\left(N_p^\alpha e_\mu C_{\eta\alpha}^{\mu}\dot{x}^k + M_{p\eta}^k + M_{p\gamma}^{kj} C_{\eta\mu}^{\gamma} A_j^\mu\right)\lambda_\varepsilon^\eta = 0$$

Due to (0.30), we deduce from this equation our <u>second invariance identity</u>:

$$\boxed{N_p^\alpha e_\mu C_{\eta\alpha}^{\mu}\dot{x}^k + M_{p\eta}^k - M_{p\gamma}^{kj} C_{\mu\eta}^{\gamma} A_j^\mu = 0} \tag{4.5}$$

4.3 The Third Invariance Identity

The final set of equations in (4.3) states that

$$N_p^\gamma \frac{\partial f_\gamma}{\partial u^\varepsilon} + M_p^\gamma \frac{\partial e_\gamma}{\partial u^\varepsilon} + M_{p\gamma}^j \frac{\partial A_j^\gamma}{\partial u^\varepsilon} + M_{p\gamma}^{jk} \frac{\partial A_{j,k}^\gamma}{\partial u^\varepsilon} = 0 \tag{4.6}$$

Let us consider this equation term by term. The first term in this equation, in view of (1.4), is

$$N_p^\gamma \frac{\partial f_\gamma}{\partial u^\varepsilon} = N_p^\gamma\left(\frac{\partial \widehat{G}_\gamma^\mu}{\partial u^\varepsilon}\mathring{f}_\mu - \mathring{e}_\mu C_{\sigma\gamma}^\alpha u_k^\eta \dot{x}^k\left[\frac{\partial \widehat{G}_\alpha^\mu}{\partial u^\varepsilon}\lambda_\eta^\sigma + \widehat{G}_\alpha^\mu \frac{\partial \lambda_\eta^\sigma}{\partial u^\varepsilon}\right]\right)$$

or by (0.35)

$$= N_p^\gamma C_{\sigma\gamma}^\alpha\left(-\widehat{G}_\alpha^\mu \lambda_\varepsilon^\sigma \mathring{f}_\mu - \mathring{e}_\mu u_k^\eta \dot{x}^k\left[\frac{\partial \widehat{G}_\alpha^\mu}{\partial u^\varepsilon}\lambda_\eta^\sigma + \widehat{G}_\alpha^\mu \frac{\partial \lambda_\eta^\sigma}{\partial u^\varepsilon}\right]\right)$$

$$= N_p^\gamma C_{\sigma\gamma}^\alpha\left(-\widehat{G}_\alpha^\mu \lambda_\varepsilon^\sigma \mathring{f}_\mu - \mathring{e}_\mu u_k^\eta \dot{x}^k\left[-\widehat{G}_\beta^\mu C_{\kappa\alpha}^\beta \lambda_\varepsilon^\kappa \lambda_\eta^\sigma + \widehat{G}_\alpha^\mu \frac{\partial \lambda_\eta^\sigma}{\partial u^\varepsilon}\right]\right)$$

By substitution from (1.4), this is

$$N_p^\gamma \frac{\partial f_\gamma}{\partial u^\varepsilon} = N_p^\gamma C_{\sigma\gamma}^\alpha\left(-\lambda_\varepsilon^\sigma\left[f_\alpha + e_\eta C_{\kappa\alpha}^\eta \lambda_\mu^\kappa u_k^\mu \dot{x}^k\right] - \mathring{e}_\mu u_k^\eta \dot{x}^k\left[-\widehat{G}_\beta^\mu C_{\kappa\alpha}^\beta \lambda_\varepsilon^\kappa \lambda_\eta^\sigma + \widehat{G}_\alpha^\mu \frac{\partial \lambda_\eta^\sigma}{\partial u^\varepsilon}\right]\right) \tag{4.7}$$

Part of this expression - namely

$$N_p^\gamma C_{\sigma\gamma}^\alpha\left(-\lambda_\varepsilon^\sigma e_\eta C_{\kappa\alpha}^\eta \lambda_\mu^\kappa u_k^\mu \dot{x}^k + \mathring{e}_\mu u_k^\eta \dot{x}^k \widehat{G}_\beta^\mu C_{\kappa\alpha}^\beta \lambda_\varepsilon^\kappa \lambda_\eta^\sigma\right)$$

may be written as

$$N_p^\gamma \lambda_\eta^\sigma u_k^\eta \dot{x}^k\left(-\lambda_\varepsilon^\beta e_\xi C_{\beta\gamma}^\alpha C_{\sigma\alpha}^\xi + e_\beta C_{\sigma\gamma}^\alpha C_{\kappa\alpha}^\beta \lambda_\varepsilon^\kappa\right)$$

that is,

$$N_p^\gamma \lambda_\eta^\sigma u_k^\eta \dot{x}^k e_\xi \lambda_\varepsilon^\beta\left(-C_{\beta\gamma}^\alpha C_{\sigma\alpha}^\xi + C_{\sigma\gamma}^\alpha C_{\beta\alpha}^\xi\right)$$

Application of (0.24) and (0.25) reduces this to

$$N_p^\gamma \lambda_\eta^\sigma u_k^\eta \dot{x}^k e_\xi \lambda_\varepsilon^\beta C_{\sigma\beta}^\alpha C_{\gamma\alpha}^\xi$$

and as a result, (4.7) is

$$N_p^\gamma \frac{\partial f_\gamma}{\partial u^\varepsilon} = N_p^\gamma C_{\sigma\gamma}^{\alpha}\left(-\lambda_\varepsilon^\sigma f_\alpha - \mathring{e}_\mu u_k^\eta \dot{x}^k \widehat{G}_\alpha^\mu \frac{\partial \lambda_\eta^\sigma}{\partial u^\varepsilon}\right) + N_p^\gamma C_{\sigma\beta}^{\alpha} \lambda_\eta^\sigma u_k^\eta \dot{x}^k e_\xi \lambda_\varepsilon^\beta C_{\gamma\alpha}^{\xi}$$

$$N_p^\gamma \frac{\partial f_\gamma}{\partial u^\varepsilon} = N_p^\gamma C_{\sigma\gamma}^{\alpha}\left(-\lambda_\varepsilon^\sigma f_\alpha - e_\alpha u_k^\eta \dot{x}^k \frac{\partial \lambda_\eta^\sigma}{\partial u^\varepsilon}\right) + N_p^\gamma \left(\frac{\partial \lambda_\varepsilon^\alpha}{\partial u^\eta} - \frac{\partial \lambda_\eta^\alpha}{\partial u^\varepsilon}\right) u_k^\eta \dot{x}^k e_\xi C_{\gamma\alpha}^{\xi}$$

or

$$N_p^\gamma \frac{\partial f_\gamma}{\partial u^\varepsilon} = -N_p^\gamma C_{\sigma\gamma}^{\alpha} \lambda_\varepsilon^\sigma f_\alpha + N_p^\gamma \frac{\partial \lambda_\varepsilon^\alpha}{\partial u^\eta} u_k^\eta \dot{x}^k e_\xi C_{\gamma\alpha}^{\xi} \tag{4.8}$$

This is the first term of (4.6).

Moreover, (1.1) and (0.37) imply that the second term in (4.6) is

$$M_p^\gamma \frac{\partial \widehat{G}_\gamma^{\beta} \mathring{e}_\beta}{\partial u^\varepsilon} = -M_p^\gamma \mathring{e}_\beta \widehat{G}_\sigma^\beta C_{\alpha\gamma}^{\sigma} \lambda_\varepsilon^\alpha$$

$$= M_p^\gamma e_\sigma C_{\gamma\alpha}^{\sigma} \lambda_\varepsilon^\alpha \tag{4.9}$$

The third term in (4.6), by application of (0.35), is

$$\frac{\partial (G_\eta^\gamma \mathring{A}_j^\eta - \lambda_\eta^\gamma u_j^\eta)}{\partial u^\varepsilon} = M_{p\gamma}^j \left[C_{\alpha\beta}^{\gamma} \lambda_\varepsilon^\alpha G_\eta^\beta \mathring{A}_j^\eta - u_j^\eta \frac{\partial \lambda_\eta^\gamma}{\partial u^\varepsilon}\right]$$

or, due to (0.46)

$$= M_{p\gamma}^j \left[C_{\alpha\beta}^{\gamma} \lambda_\varepsilon^\alpha A_j^\beta + C_{\alpha\beta}^{\gamma} \lambda_\varepsilon^\alpha \lambda_\eta^\beta u_j^\eta - u_j^\eta \frac{\partial \lambda_\eta^\gamma}{\partial u^\varepsilon}\right]$$

and by (0.36)

$$= M_{p\gamma}^j C_{\alpha\beta}^{\gamma} \lambda_\varepsilon^\alpha A_j^\beta - M_{p\gamma}^j \frac{\partial \lambda_\varepsilon^\gamma}{\partial u^\eta} u_j^\eta \tag{4.10}$$

The fourth term in (4.6), due to (0.50), is

$$M^{jk}_{p\gamma}\frac{\partial A^{\gamma}_{j,k}}{\partial u^{\varepsilon}} = M^{jk}_{p\gamma}\left[\frac{\partial G^{\gamma}_{\eta}}{\partial u^{\varepsilon}}\mathring{A}^{\eta}_{j,k} + \frac{\partial}{\partial u^{\varepsilon}}\left(\frac{\partial G^{\gamma}_{\eta}}{\partial u^{\beta}}\mathring{A}^{\eta}_{j} - \frac{\partial \lambda^{\gamma}_{\eta}}{\partial u^{\beta}}u^{\eta}_{j}\right)u^{\beta}_{k} - \frac{\partial \lambda^{\gamma}_{\eta}}{\partial u^{\varepsilon}}u^{\eta}_{j\,k}\right]$$

Since by (4.4) and (0.42) $M^{jk}_{p\gamma}$ and $u^{\beta}_{j\,k}$ are skew-symmetric and symmetric respectively in j and k, this is

$$M^{jk}_{p\gamma}\left[\frac{\partial G^{\gamma}_{\eta}}{\partial u^{\varepsilon}}\mathring{A}^{\eta}_{j,k} + \left(\frac{\partial}{\partial u^{\varepsilon}}\left[\frac{\partial G^{\gamma}_{\eta}}{\partial u^{\beta}}\right]\mathring{A}^{\eta}_{j} - \frac{\partial}{\partial u^{\varepsilon}}\left[\frac{\partial \lambda^{\gamma}_{\eta}}{\partial u^{\beta}}\right]u^{\eta}_{j}\right)u^{\beta}_{k}\right]$$

Equations (0.35) and (0.36) allow this to be expressed as

$$M^{jk}_{p\gamma}\left[C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}G^{\nu}_{\eta}\mathring{A}^{\eta}_{j,k} + \left(\frac{\partial}{\partial u^{\beta}}\left[C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}G^{\nu}_{\eta}\right]\mathring{A}^{\eta}_{j} - \frac{\partial}{\partial u^{\beta}}\left[C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}\lambda^{\nu}_{\eta} + \frac{\partial \lambda^{\gamma}_{\varepsilon}}{\partial u^{\eta}}\right]u^{\eta}_{j}\right)u^{\beta}_{k}\right]$$

Due to the skew-symmetry (4.4) of $M^{jk}_{p\gamma}$ in its upper indices, and the symmetry in j and k of the term involving the second derivative of $\lambda^{\gamma}_{\varepsilon}$, the term involving the second derivative of $\lambda^{\gamma}_{\varepsilon}$ vanishes. This leaves

$$M^{jk}_{p\gamma}\left[C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}G^{\nu}_{\eta}\mathring{A}^{\eta}_{j,k} + \left(\frac{\partial}{\partial u^{\beta}}\left[C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}G^{\nu}_{\eta}\right]\mathring{A}^{\eta}_{j} - \frac{\partial(\lambda^{\alpha}_{\varepsilon}\lambda^{\nu}_{\eta})}{\partial u^{\beta}}C^{\gamma}_{\alpha\nu}u^{\eta}_{j}\right)u^{\beta}_{k}\right]$$

which is

$$M^{jk}_{p\gamma}C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}G^{\nu}_{\eta}\mathring{A}^{\eta}_{j,k} +$$
$$M^{jk}_{p\gamma}\left(\frac{\partial \lambda^{\alpha}_{\varepsilon}}{\partial u^{\beta}}C^{\gamma}_{\alpha\nu}G^{\nu}_{\eta}\mathring{A}^{\eta}_{j} + \frac{\partial G^{\nu}_{\eta}}{\partial u^{\beta}}C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}\mathring{A}^{\eta}_{j} - \left[\frac{\partial \lambda^{\alpha}_{\varepsilon}}{\partial u^{\beta}}C^{\gamma}_{\alpha\nu}\lambda^{\nu}_{\eta} + \frac{\partial \lambda^{\nu}_{\eta}}{\partial u^{\beta}}C^{\gamma}_{\alpha\nu}\lambda^{\alpha}_{\varepsilon}\right]u^{\eta}_{j}\right)u^{\beta}_{k}$$

In abbreviated notation this is

$$M^{jk}_{p\gamma}C^{\gamma}_{\alpha\nu}\left(\lambda^{\alpha}_{\varepsilon}G^{\nu}_{\eta}\mathring{A}^{\eta}_{j,k} + \lambda^{\alpha}_{\varepsilon,k}G^{\nu}_{\eta}\mathring{A}^{\eta}_{j} + G^{\nu}_{\eta,k}\lambda^{\alpha}_{\varepsilon}\mathring{A}^{\eta}_{j} - \left[\lambda^{\alpha}_{\varepsilon,k}\lambda^{\nu}_{\eta} + \lambda^{\nu}_{\eta,k}\lambda^{\alpha}_{\varepsilon}\right]u^{\eta}_{j}\right)$$

or

$$M^{jk}_{p\gamma}C^{\gamma}_{\alpha\nu}\left(\lambda^{\alpha}_{\varepsilon,k}\left[G^{\nu}_{\eta}\mathring{A}^{\eta}_{j} - \lambda^{\nu}_{\eta}u^{\eta}_{j}\right] + \lambda^{\alpha}_{\varepsilon}\left[G^{\nu}_{\eta}\mathring{A}^{\eta}_{j,k} + G^{\nu}_{\eta,k}\mathring{A}^{\eta}_{j} - \lambda^{\nu}_{\eta,k}u^{\eta}_{j}\right]\right)$$

Next, we can simplify this using (0.46) and also insert a term which vanishes (due to symmetry considerations from (4.4)). As a result, it is found that

$$M^{jk}_{p\gamma}C^{\gamma}_{\alpha\nu}\left(\lambda^{\alpha}_{\varepsilon,k}A^{\nu}_{j} + \lambda^{\alpha}_{\varepsilon}\left[G^{\nu}_{\eta}\mathring{A}^{\eta}_{j,k} + G^{\nu}_{\eta,k}\mathring{A}^{\eta}_{j} - \lambda^{\nu}_{\eta,k}u^{\eta}_{j} - \lambda^{\alpha}_{\eta}u^{\eta}_{j\,k}\right]\right)$$

and (0.50) can be applied to obtain

$$M^{jk}_{p\gamma}C^{\gamma}_{\alpha\nu}\left(\lambda^{\alpha}_{\varepsilon,k}A^{\nu}_{j} + \lambda^{\alpha}_{\varepsilon}A^{\nu}_{j,k}\right)$$

Equation (4.5) then shows that

$$\left(M^{k}_{p\alpha} + N^{\eta}_{p}e_{\mu}C^{\mu}_{\alpha\eta}\dot{x}^{k}\right)\lambda^{\alpha}_{\varepsilon,k} + \lambda^{\alpha}_{\varepsilon}M^{jk}_{p\gamma}C^{\gamma}_{\alpha\nu}A^{\nu}_{j,k}$$

Therefore,

$$M^{jk}_{p\gamma}\frac{\partial A^{\gamma}_{j,k}}{\partial u^{\varepsilon}} = \left(M^{k}_{p\alpha} + N^{\eta}_{p}e_{\mu}C^{\mu}_{\alpha\eta}\dot{x}^{k}\right)\lambda^{\alpha}_{\varepsilon,k} + \lambda^{\alpha}_{\varepsilon}M^{jk}_{p\gamma}C^{\gamma}_{\alpha\nu}A^{\nu}_{j,k} \tag{4.11}$$

Thus, we have expressions (4.8), (4.9), (4.10) and (4.11) for the terms in (4.6). Substitution of these in the latter yields the following form of the third invariance identity:

$$-N_p^{\gamma} C_{\sigma\gamma}^{\alpha} \lambda_\varepsilon^{\sigma} f_\alpha + N_p^{\gamma} \frac{\partial \lambda_\varepsilon^{\alpha}}{\partial u^\eta} u_k^\eta \dot{x}^k e_\xi C_{\gamma\alpha}^{\xi} + M_p^{\gamma} e_\sigma C_{\gamma\alpha}^{\sigma} \lambda_\varepsilon^{\alpha} + M_{p\gamma}^{j} C_{\alpha\beta}^{\gamma} \lambda_\varepsilon^{\alpha} A_j^{\beta}$$

$$- M_{p\gamma}^{k} \lambda_{\varepsilon,k}^{\gamma} + \left(M_{p\alpha}^{k} + N_p^{\eta} e_\mu C_{\alpha\eta}^{\mu} \dot{x}^k\right) \lambda_{\varepsilon,k}^{\alpha} + \lambda_\varepsilon^{\alpha} M_{p\gamma}^{jk} C_{\alpha\nu}^{\gamma} A_{j,k}^{\nu} = 0$$

Some terms evidently cancel, and the resulting equation is

$$-N_p^{\gamma} C_{\sigma\gamma}^{\alpha} \lambda_\varepsilon^{\sigma} f_\alpha + M_p^{\gamma} e_\sigma C_{\gamma\alpha}^{\sigma} \lambda_\varepsilon^{\alpha} + M_{p\gamma}^{j} C_{\alpha\beta}^{\gamma} \lambda_\varepsilon^{\alpha} A_j^{\beta} + \lambda_\varepsilon^{\alpha} M_{p\gamma}^{jk} C_{\alpha\nu}^{\gamma} A_{j,k}^{\nu} = 0$$

or, since by (0.30), $\det(\lambda_\varepsilon^{\alpha}) \neq 0$,

$$N_p^{\gamma} C_{\gamma\alpha}^{\sigma} f_\sigma + M_p^{\gamma} e_\sigma C_{\gamma\alpha}^{\sigma} + M_{p\gamma}^{j} C_{\alpha\beta}^{\gamma} A_j^{\beta} + M_{p\gamma}^{jk} C_{\alpha\nu}^{\gamma} A_{j,k}^{\nu} = 0$$

By (4.4) and the definition (0.51) of the field strengths F_{jk}^{ε}, this becomes

$$N_p^{\gamma} C_{\gamma\alpha}^{\sigma} f_\sigma + M_p^{\gamma} e_\sigma C_{\gamma\alpha}^{\sigma} + M_{p\gamma}^{j} C_{\alpha\beta}^{\gamma} A_j^{\beta} + \frac{1}{2} M_{p\gamma}^{jk} C_{\alpha\nu}^{\gamma} \left(F_{jk}^{\nu} - C_{\sigma\mu}^{\nu} A_k^{\sigma} A_j^{\mu}\right) = 0$$

or, due to (4.4) and (4.5)

$$\boxed{N_p^{\gamma} C_{\gamma\alpha}^{\sigma} f_\sigma + M_p^{\gamma} e_\sigma C_{\gamma\alpha}^{\sigma} - N_p^{\sigma} e_\mu C_{\gamma\sigma}^{\mu} C_{\alpha\beta}^{\gamma} A_j^{\beta} \dot{x}^j + \frac{1}{2} M_{p\gamma}^{jk} C_{\alpha\nu}^{\gamma} F_{jk}^{\nu} = 0} \qquad (4.12)$$

This is the third invariance identity:

In retrospect, we have in this chapter established the following fact.[16]

---

[16] This fact is analogous to a result obtained by Rund [13], p. 113; however, the third invariance identity has been expressed in a different form and some arguments which are not affected by gauge transformations have been included in M_p.

Fact: *Suppose that M_p denotes a set of C^1 functions of the form (4.1), whose arguments $e_\alpha, f_\alpha, A_j^\alpha, A_{j,k}^\alpha$ transform under the action of an r-dimensional Lie group according to (1.1), (1.4), (0.46) and (0.50) respectively (whose arguments $\dot{x}^i$, $\ddot{x}^i$ g_{ij}, $g_{ij,k}$ are unaffected by gauge transformations). Equations (4.4), (4.5) and (4.12) constitute a set of necessary and sufficient conditions in order that the functions M_p be gauge-invariant.*

CHAPTER 5

THE REQUIRED FORM OF THE EQUATIONS OF MOTION

In this chapter, the invariance identities from chapter 4 are applied to the equations of motion from chapter 3. The result is a remarkably simple form for both the function L (which determines our Lagrangian) and for the vector E_k (which determines our equations of motion). Some very important intermediate results are obtained during the analysis.

It is demonstrated that a general gauge-invariant function M_p of the form (4.1) must satisfy the functional relationship

$$M_p = M_p'(e_\alpha , \frac{De_\alpha}{D\tau} , F_{j\ k}^{\ \alpha} , \dot{x}^j , \ddot{x}^j , g_{jk} , g_{jk,h})$$

for some C^1 function M_p'.

When this fact (together with previous equations) is applied to the form (3.6) of the vector E_k, it is deduced that the function L must be expressible as

$$L = C\sqrt{|\gamma|} + G_\alpha(e_\eta)A_j^\alpha \dot{x}^j$$

for some C^3 scalar functions G_α ($\alpha = 1,\ldots,r$) and nonzero constant C. The functions G_α are found to comprise the components of a type (0,1) adjoint object whose adjoint absolute derivative vanishes. If the signature of the metric is chosen such that $\gamma > 0$ for all time-like vectors, then the resulting gauge-invariant vector E_k is expressible using proper time, in the form

$$E_k = C\, g_{ik}\frac{D\dot{x}^i}{D\tau} - G_\alpha F_{j\ k}^{\ \alpha}\dot{x}^j$$

so that the motion of the particle is governed by the equation

$$g_{ik}\frac{D\dot{x}^i}{D\tau} = \frac{G_\alpha}{C} F_{j\ k}^{\ \alpha} \dot{x}^j$$

5.1 The Invariance Identities and General Gauge-Invariant Functions

We begin by obtaining a functional equation that functions M_p of the form (4.1) must satisfy as a consequence of the first and second invariance identities (equations (4.4) and (4.5)). To save writing we let "..." denote the arguments "$\dot{x}^j$, $\ddot{x}^j$, g_{jk}, $g_{jk,h}$" of (4.1) which do not depend upon u^α, u_j^α, or $u_{j\ k}^{\ \alpha}$, i.e.

$$M_p = M_p(e_\alpha, f_\alpha, A_j^\alpha, A_{j,k}^\alpha, \ldots)$$

The invariance identity (4.4) is valid if and only if each function M_p depends skew-symmetrically through j and k upon the quantities $A_{j,k}^{\alpha}$. Thus M_p must be expressible as

$$M_p(e_\alpha, f_\alpha, A_j^\alpha, \frac{1}{2}(A_{j,k}^\alpha - A_{j,k}^\alpha), \ldots)$$

This however, may be written as

$$M_p(e_\alpha, f_\alpha, A_j^\alpha, \frac{1}{2}(A_{j,k}^\alpha - A_{j,k}^\alpha + C_{\beta\gamma}^{\alpha} A_j^\beta A_k^\gamma) - \frac{1}{2} C_{\beta\gamma}^{\alpha} A_j^\beta A_k^\gamma, \ldots)$$

or, due to (0.51), as

$$M_p(e_\alpha, f_\alpha, A_j^\alpha, \frac{1}{2} F_{j\ k}^{\ \alpha} - \frac{1}{2} C_{\beta\gamma}^{\alpha} A_j^\beta A_k^\gamma, \ldots)$$

or

$$M_p(e_\alpha\,, f_\alpha - e_\eta C^{\,\eta}_{\beta\alpha} A^\beta_k \dot{x}^k + e_\eta C^{\,\eta}_{\beta\alpha} A^\beta_k \dot{x}^k\,,\, A^\alpha_j\,,\, \tfrac{1}{2} F_j{}^\alpha_k + \tfrac{1}{2} C^{\,\alpha}_{\beta\gamma} A^\beta_j A^\gamma_k\,,\ldots)$$

Because of (1.5), this is

$$M_p(e_\alpha\,, \frac{De_\alpha}{D\tau} + e_\eta C^{\,\eta}_{\beta\alpha} A^\beta_k \dot{x}^k\,,\, A^\alpha_j\,,\, \tfrac{1}{2} F_j{}^\alpha_k + \tfrac{1}{2} C^{\,\alpha}_{\beta\gamma} A^\beta_j A^\gamma_k\,,\ldots)$$

and we deduce that the invariance identity (4.4) is equivalent to the requirement that

$$M_p(e_\alpha, f_\alpha\,, A^\alpha_j,\, A^\alpha_{j,k}\,,\ldots) = M_p(e_\alpha\,, \frac{De_\alpha}{D\tau} + e_\eta C^{\,\eta}_{\beta\alpha} A^\beta_k \dot{x}^k\,,\, A^\alpha_j\,,\, \tfrac{1}{2} F_j{}^\alpha_k + \tfrac{1}{2} C^{\,\alpha}_{\beta\gamma} A^\beta_j A^\gamma_k\,,\ldots) \tag{5.1}$$

This holds only if there exist C^1 functions M''_p ($p = 1,\ldots,n$) such that

$$M_p \;=\; M''_p(e_\alpha\,, F_\alpha\,, A^\alpha_j,\, F_j{}^\alpha_k\,, \dot{x}^j\,, \ddot{x}^j\,,\, g_{jk}\,,\, g_{jk,h}\,) \tag{5.2}$$

where we have introduced

$$F_\alpha = \frac{De_\alpha}{D\tau} = f_\alpha - e_\eta C^{\,\eta}_{\beta\alpha} A^\beta_k \dot{x}^k \tag{5.3}$$

Let us now consider the invariance identity (4.5). We suppose that the first invariance identity holds, and express (4.5) in terms of the functions M''_p of (5.2). Equations (5.1) and (5.2) require that

$$M''_p(e_\alpha, F_\alpha, A^\alpha_j, F_j{}^\alpha_k,\ldots) \;=\; M_p(e_\alpha, F_\alpha + e_\eta C^{\,\eta}_{\beta\alpha} A^\beta_k \dot{x}^k, A^\alpha_j, \tfrac{1}{2} F_j{}^\alpha_k + \tfrac{1}{2} C^{\,\alpha}_{\beta\gamma} A^\beta_j A^\gamma_k,\ldots) \tag{5.4}$$

and partial differentiation of (5.4) with respect to A^σ_q then shows that

$$\frac{\partial M''_p}{\partial A^\sigma_q} \;=\; \frac{\partial\, M_p(e_\alpha,\, F_\alpha + e_\eta C^{\,\eta}_{\beta\alpha} A^\beta_k \dot{x}^k,\, A^\alpha_j,\, \tfrac{1}{2} F_j{}^\alpha_k + \tfrac{1}{2} C^{\,\alpha}_{\beta\gamma} A^\beta_j A^\gamma_k\,,\ldots)}{\partial A^\sigma_q}$$

$$= N_p^\gamma e_\eta C_{\beta\gamma}^{\eta} \delta_\sigma^\beta \delta_k^q \dot{x}^k + M_{p\sigma}^q + \frac{1}{2} M_{p\alpha}^{jk} C_{\beta\gamma}^{\alpha} \left(\delta_\sigma^\beta \delta_j^q A_k^\gamma + A_j^\beta \delta_\sigma^\gamma \delta_k^q \right)$$

$$= N_p^\gamma e_\eta C_{\sigma\gamma}^{\eta} \dot{x}^q + M_{p\sigma}^q + M_{p\alpha}^{qk} C_{\sigma\gamma}^{\alpha} A_k^\gamma$$

But (4.5) states that the right hand side of this equation is zero. Hence, the invariance identity (4.5) in terms of $M_p^{''}$ is

$$\boxed{\frac{\partial M_p^{''}}{\partial A_q^\sigma} = 0} \tag{5.5}$$

which holds only if there exist n functions $M_p^{'}$ in terms of which

$$\boxed{M_p = M_p^{'}(e_\alpha , F_\alpha , F_{j\,k}^{\alpha} , \dot{x}^j, \ddot{x}^j, g_{jk}, g_{jk,h})} \tag{5.6}$$

Thus, *the invariance identities (4.4) and (4.5) imply that there exist n C^1 functions $M_p^{'}$ for which (5.6) is valid* (where the quantities F_α , $F_{j\,k}^{\alpha}$ in (5.6) are defined by (5.3) and (0.51) respectively). If the arguments of M_p were all independent this statement would have a converse - however, it shall be shown that they are not all independent.

5.2 The Required Form of the Lagrangian

We now apply these results to the vector E_k of (3.6), and obtain in consequence, a simple functional expression for L and its associated Lagrangian $\overset{\circ}{L}$. In equation (3.6) it is *not* assumed that τ denotes proper time, so the quantities x^h, $\dot{x}^h$, $\ddot{x}^h$, u^α, u_h^α, $u_{h\,j}^{\alpha}$ in that equation are independent (except for the symmetry of u_{hj}^{α}). As a result, E_k can be gauge-invariant only if its partial derivatives

$$\frac{\partial E_k}{\partial \ddot{x}^j} = \left(A_k^{\alpha}\frac{\partial^2 L^*}{\partial\gamma\partial A^{\alpha}} + 2\frac{\partial^2 L^*}{\partial\gamma\partial\gamma}\,g_{ik}\dot{x}^i\right)2\,g_{jm}\dot{x}^m + 2\frac{\partial L^*}{\partial\gamma}g_{jk} \tag{5.7}$$

are also gauge-invariant. It is by use of this crucial fact and the invariance identities, that we obtain the desired form for L.

The right-hand-side of (5.7) is of the form (4.1), and so can be gauge invariant only if it satisfies (5.5) - namely,

$$\frac{\partial}{\partial A_h^{\mu}}\left[\left(A_k^{\alpha}\frac{\partial^2 L^*}{\partial\gamma\partial A^{\alpha}} + 2\frac{\partial^2 L^*}{\partial\gamma\partial\gamma}\,g_{ik}\dot{x}^i\right)2\,g_{jm}\dot{x}^m + 2\frac{\partial L^*}{\partial\gamma}g_{jk}\right] = 0$$

that is

$$\left(\delta_k^h\,\delta_{\mu}^{\alpha}\frac{\partial^2 L^*}{\partial\gamma\partial A^{\alpha}} + A_k^{\alpha}\frac{\partial^3 L^*}{\partial A^{\eta}\partial\gamma\partial A^{\alpha}}\frac{\partial A^{\eta}}{\partial A_h^{\mu}} + 2\frac{\partial^3 L^*}{\partial A^{\eta}\partial\gamma\partial\gamma}\frac{\partial A^{\eta}}{\partial A_h^{\mu}}\,g_{ik}\dot{x}^i\right)2\,g_{jm}\dot{x}^m$$
$$+\,2\frac{\partial^2 L^*}{\partial A^{\eta}\partial\gamma}\frac{\partial A^{\eta}}{\partial A_h^{\mu}}g_{jk} = 0$$

or

$$\left(\frac{\partial^2 L^*}{\partial\gamma\partial A^{\mu}}\delta_k^h + \frac{\partial^3 L^*}{\partial A^{\mu}\partial\gamma\partial A^{\alpha}}A_k^{\alpha}\dot{x}^h + 2\frac{\partial^3 L^*}{\partial A^{\mu}\partial\gamma\partial\gamma}g_{ik}\dot{x}^i\dot{x}^h\right)2\,g_{jm}\dot{x}^m + 2\frac{\partial^2 L^*}{\partial A^{\mu}\partial\gamma}g_{jk}\dot{x}^h = 0 \tag{5.8}$$

By putting $k = h'$ (for any index $h' \neq h$) in (5.8) we obtain

$$\left[\left(\frac{\partial^3 L^*}{\partial A^{\mu}\partial\gamma\partial A^{\alpha}}A_{h'}^{\alpha} + 2\frac{\partial^3 L^*}{\partial A^{\mu}\partial\gamma\partial\gamma}g_{ih'}\dot{x}^i\right)2\,g_{jm}\dot{x}^m + 2\frac{\partial^2 L^*}{\partial A^{\mu}\partial\gamma}g_{jh'}\right]\dot{x}^h = 0$$

According to (1.8) at least one component of the tangent vector is nonzero, so if we let $\dot{x}^h$ be such a component we find that

$$\left(\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma \partial A^\alpha} A^\alpha_{h'} + 2\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma\partial\gamma} g_{ih'}\dot{x}^i\right) 2 g_{jm}\dot{x}^m + 2\frac{\partial^2 L^*}{\partial A^\mu \partial\gamma} g_{jh'} = 0 \qquad \text{(for any } h' \neq h\text{)}$$

If we multiply this equation by g^{hj} and sum on j, we obtain

$$\left(\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma \partial A^\alpha} A^\alpha_{h'} + 2\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma\partial\gamma} g_{ih'}\dot{x}^i\right) 2\dot{x}^h = 0$$

or, since h is such that $\dot{x}^h \neq 0$,

$$\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma \partial A^\alpha} A^\alpha_{h'} + 2\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma\partial\gamma} g_{ih'}\dot{x}^i = 0 \qquad (5.9)$$

Now set $k = h' \neq h$ in (5.8). Then,

$$\left(\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma \partial A^\alpha} A^\alpha_{h}\dot{x}^h + 2\frac{\partial^3 L^*}{\partial A^\mu \partial\gamma\partial\gamma} g_{ih'}\dot{x}^i\dot{x}^h\right) 2 g_{jm}\dot{x}^m + 2\frac{\partial^2 L^*}{\partial A^\mu \partial\gamma} g_{jh'}\dot{x}^h = 0$$

or after substitution from (5.9),

$$2\frac{\partial^2 L^*}{\partial A^\mu \partial\gamma} g_{jh'}\dot{x}^h = 0$$

which, since $\det(g_{jk}) \neq 0$ and $\dot{x}^h \neq 0$, implies that

$$\frac{\partial^2 L^*}{\partial A^\mu \partial\gamma} = 0 \qquad (5.10)$$

Equations (5.10) and (2.32) together require that there exists a pair of C^3 functions L_I, L_{II} such that

$$L^* = L_I(e_\alpha, \gamma) + L_{II}(e_\alpha, A^\alpha) \tag{5.11}$$

Since their arguments are scalars, these functions must be scalars.

This result can be made even stronger due to (3.4). Substitution of (5.10) into (3.4) shows that

$$\frac{\partial^2 L_{II}}{\partial A^\beta \partial A^\eta} A_p^\beta = 0$$

and partial differentiation with respect to the independent parameters u_h^μ then yields

$$-\frac{\partial^2 L_{II}}{\partial A^\beta \partial A^\eta} \lambda_\mu^\beta \delta_p^h - \frac{\partial^3 L_{II}}{\partial A^\varepsilon \partial A^\beta \partial A^\eta} A_p^\beta \lambda_\mu^\varepsilon \dot{x}^h = 0$$

or, by (0.30)

$$-\frac{\partial^2 L_{II}}{\partial A^\alpha \partial A^\eta} \delta_p^h - \frac{\partial^3 L_{II}}{\partial A^\alpha \partial A^\beta \partial A^\eta} A_p^\beta \dot{x}^h = 0 \tag{5.12}$$

But since L_{II} and its arguments are scalars, the partial derivatives shown are scalars. Hence, each term in (5.12) (due to (0.14) and (0.47)) is a type (1,1) tensor. Under an admissible coordinate transformation, we therefore have

$$-\frac{\partial^2 \overline{L}_{II}}{\partial \overline{A}^\alpha \partial \overline{A}^\eta} \delta_p^h - \frac{\partial^3 \overline{L}_{II}}{\partial \overline{A}^\alpha \partial \overline{A}^\beta \partial \overline{A}^\eta} \overline{A}_p^\beta \dot{\overline{x}}^h = 0$$

so that

$$-\frac{\partial^2 L_{II}}{\partial A^\alpha \partial A^\eta} \delta_p^h - \frac{\partial^3 L_{II}}{\partial A^\alpha \partial A^\beta \partial A^\eta} B_p^j A_j^\beta \widehat{B}_i^h \dot{x}^i = 0$$

Differentiation this equation with respect to B_k^q and use of (0.3) shows that

$$\frac{\partial^3 L_{II}}{\partial A^\alpha \partial A^\beta \partial A^\eta}\left(\delta_q^j\, \delta_p^k\, \widehat{B}_i^h - B_p^j\, \widehat{B}_q^h\, \widehat{B}_i^k\right)A_j^\beta \dot{x}^i = 0$$

This equation may be evaluated using the identity transformation ($B_j^h = \delta_j^h$, $\widehat{B}_j^h = \delta_j^h$) to obtain

$$\frac{\partial^3 L_{II}}{\partial A^\alpha \partial A^\beta \partial A^\eta}\left(\delta_q^j\, \delta_p^k\, \delta_i^h - \delta_p^j\, \delta_q^h\, \delta_i^k\right)A_j^\beta \dot{x}^i = 0$$

or

$$\frac{\partial^3 L_{II}}{\partial A^\alpha \partial A^\beta \partial A^\eta}\left(\delta_p^k\, A_q^\beta \dot{x}^h - \delta_q^h\, A_p^\beta \dot{x}^k\right) = 0$$

Since $n > 1$, we may set $p = k \neq q$ to obtain

$$\frac{\partial^3 L_{II}}{\partial A^\alpha \partial A^\beta \partial A^\eta} A_q^\beta \dot{x}^h = 0$$

which when substituted into (5.12) shows that

$$\frac{\partial^2 L_{II}}{\partial A^\epsilon \partial A^\eta} = 0 \tag{5.13}$$

Equations (5.13) and (5.11) imply that L^* must be such that

$$L^* = L_I(e_\eta, \gamma) + G_\alpha(e_\eta) A^\alpha \tag{5.14}$$

for some C^3 functions L_I and G_α.

Because $L = L^*$ is positively homogeneous of first degree in the components $\dot{x}^h$ it follows from (5.14) that

$$\left(L_I(e_\eta, \gamma) + G_\alpha(e_\eta)A_j^\alpha \dot{x}^j\right)t = L_I(e_\eta, t^2\gamma) + G_\alpha(e_\eta)A_j^\alpha t\,\dot{x}^j$$

for arbitrary $t > 0$, and hence that

$$L_I(e_\eta, \gamma)t = L_I(e_\eta, t^2\gamma)$$

Differentiation of this equation with respect to t, followed by evaluation at $t = 1$, shows that

$$L_I(e_\eta, \gamma) = 2\gamma\frac{\partial L_I(e_\eta, \gamma)}{\partial \gamma}$$

from which it is concluded that there exists a function C such that

$$L_I = \sqrt{|\gamma|}\,C(e_\eta)$$

Consequently (5.14) implies that there exist scalar functions C and G_α, such that

$$\boxed{L^* = \sqrt{|\gamma|}\,C(e_\eta) + G_\alpha(e_\eta)A^\alpha} \tag{5.15}$$

By putting (5.15) into (3.6), and assuming that the signature of the matrix $[g_{ij}]$ is chosen such that $\gamma > 0$ for all time-like vectors we obtain

$$E_p = \frac{C\,g_{ip}}{\sqrt{\gamma}}\frac{D\dot{x}^i}{D\tau} + \left(\frac{\partial G_\alpha}{\partial e_\sigma}A_p^\alpha + \frac{\partial C}{\partial e_\sigma}\frac{g_{ip}\dot{x}^i}{\sqrt{\gamma}}\right)f_\sigma - G_\alpha\left(A_{j,p}^\alpha - A_{p,j}^\alpha\right)\dot{x}^j - \frac{C\,g_{ip}\dot{x}^i}{2\gamma^{3/2}}\frac{d\gamma}{d\tau} \tag{5.16}$$

This expression will be greatly simplified in section (5.4), after we have uncovered some properties of the functions C and G_α.

5.3 The Role of the Coupling Parameters in the Lagrangian

The equations of sections 5.1 and 5.2 are applied in this section to deduce some important properties of the unknown functions C and G_α in (5.15). We continue to assume that $\gamma > 0$ for all time-like vectors.

First consider the function C. Because L_I is a C^3 scalar function, *C must be a C^3 scalar function.* Substitution of (5.15) into (1.9) shows that C must also satisfy

$$C\frac{\dot{x}^i\dot{x}^j}{2\sqrt{\gamma}} \neq 0$$

Therefore, if j is chosen such that $\dot{x}^j \neq 0$ and i is set equal to j, it follows that $C \neq 0$ (i.e. *C is either strictly negative or strictly positive on the domain of L*). Finally, according to (5.7) and (5.10), the partial derivatives

$$\frac{\partial E_k}{\partial \ddot{x}^j} = 4\frac{\partial^2 L^*}{\partial\gamma\partial\gamma} g_{ki}\dot{x}^i g_{jm}\dot{x}^m + 2\frac{\partial L^*}{\partial\gamma} g_{jk}$$

must be gauge-invariant. By substitution of L^* from (5.15) into this equation, it is found that (since $\gamma > 0$)

$$\frac{\partial E_k}{\partial \ddot{x}^j} = C\left[-\frac{g_{ki}\dot{x}^i g_{jm}\dot{x}^m}{(g_{pq}\dot{x}^p\dot{x}^q)^{3/2}} + \frac{g_{jk}}{\sqrt{g_{pq}\dot{x}^p\dot{x}^q}}\right]$$

which is

$$\frac{\partial E_k}{\partial \ddot{x}^j} = C\frac{\partial^2\sqrt{g_{pq}\dot{x}^p\dot{x}^q}}{\partial\dot{x}^k\partial\dot{x}^j}$$

and therefore, either C is gauge-invariant or

$$\frac{\partial^2 \sqrt{g_{pq}\dot{x}^p\dot{x}^q}}{\partial \dot{x}^k \partial \dot{x}^j}$$

vanishes. The latter alternative is impossible on any open region of the tangent bundle (since $\sqrt{g_{pq}\dot{x}^p\dot{x}^q}$ is not linear in $\dot{x}^p$). Hence *C must be gauge-invariant* -

$$\boxed{C(e_\sigma) = C(\mathring{e}_\sigma)} \tag{5.17}$$

Differentiation of (5.17) with respect to u^α and application of (1.2), then shows that

$$\frac{\partial C}{\partial e_\sigma} e_\mu C^{\mu}_{\varepsilon\sigma} \lambda^{\varepsilon}_{\alpha} = 0$$

or

$$\boxed{e_\mu C^{\mu}_{\varepsilon\sigma} \frac{\partial C}{\partial e_\sigma} = 0} \tag{5.18}$$

Now let us consider the functions G_α. To do so, we let M_p'' denote the vector obtained by replacing f_σ with $F_\sigma + e_\beta C^{\beta}_{\mu\sigma} A^{\mu}_{j} \dot{x}^j$ and $A^{\alpha}_{j,p}$ with $\frac{1}{2}\left(F_{jp}^{\alpha} + C^{\alpha}_{\beta\varepsilon} A^{\beta}_{j} A^{\varepsilon}_{p}\right)$ in the expression (5.16) for E_p. We find that

$$M_p'' = \frac{C\, g_{ip}}{\sqrt{\gamma}} \frac{D\dot{x}^i}{D\tau} + \left(\frac{\partial G_\alpha}{\partial e_\sigma} A^{\alpha}_{p} + \frac{\partial C}{\partial e_\sigma} \frac{g_{ip}\dot{x}^i}{\sqrt{\gamma}}\right)\left(F_\sigma + e_\beta C^{\beta}_{\mu\sigma} A^{\mu}_{j} \dot{x}^j\right) - G_\alpha\left(F_{jp}^{\alpha} + C^{\alpha}_{\beta\varepsilon} A^{\beta}_{j} A^{\varepsilon}_{p}\right)\dot{x}^j$$

$$- \frac{C\, g_{ip}\, \dot{x}^i}{2\, \gamma^{3/2}} \frac{d\gamma}{d\tau}$$

Since E_p is required to be gauge-invariant, (5.5) requires that these functions satisfy

$$\frac{\partial M_p''}{\partial A_k^\eta} = 0$$

which is

$$\frac{\partial}{\partial A_k^\eta}\left[\left(\frac{\partial G_\alpha}{\partial e_\sigma} A_p^\alpha + \frac{\partial C}{\partial e_\sigma}\frac{g_{ip}\dot{x}^i}{\sqrt{\gamma}}\right)\left(F_\sigma + e_\beta C_{\mu\ \sigma}^{\ \beta} A_j^\mu \dot{x}^j\right)\right] - G_\alpha \frac{\partial}{\partial A_k^\eta}\left(C_{\beta\ \varepsilon}^{\ \alpha} A_j^\beta A_p^\varepsilon \dot{x}^j\right) = 0$$

i.e.

$$\frac{\partial G_\eta}{\partial e_\sigma}\left(F_\sigma + e_\beta C_{\mu\ \sigma}^{\ \beta} A_j^\mu \dot{x}^j\right)\delta_p^k + \left(\frac{\partial G_\alpha}{\partial e_\sigma} A_p^\alpha + \frac{\partial C}{\partial e_\sigma}\frac{g_{ip}\dot{x}^i}{\sqrt{\gamma}}\right) e_\beta C_{\eta\ \sigma}^{\ \beta} \dot{x}^k$$

$$- G_\alpha\left(C_{\eta\ \varepsilon}^{\ \alpha} A_p^\varepsilon \dot{x}^k + C_{\beta\ \eta}^{\ \alpha} A_j^\beta \dot{x}^j \delta_p^k\right) = 0$$

Because of (5.18), this equation is reduced to

$$\frac{\partial G_\eta}{\partial e_\sigma}\left(F_\sigma + e_\beta C_{\mu\ \sigma}^{\ \beta} A_j^\mu \dot{x}^j\right)\delta_p^k + \frac{\partial G_\alpha}{\partial e_\sigma} A_p^\alpha e_\beta C_{\eta\ \sigma}^{\ \beta} \dot{x}^k - G_\alpha\left(C_{\eta\ \varepsilon}^{\ \alpha} A_p^\varepsilon \dot{x}^k + C_{\beta\ \eta}^{\ \alpha} A_j^\beta \dot{x}^j \delta_p^k\right) = 0 \tag{5.19}$$

Thus we have an equation that the functions G_α must satisfy.

Equations (5.19) have some important consequences. Let us multiply (5.19) by $\dot{x}^p$ and sum on p to obtain

$$\frac{\partial G_\eta}{\partial e_\sigma}\left(F_\sigma + e_\beta C_{\mu\ \sigma}^{\ \beta} A^\mu\right)\dot{x}^k + \frac{\partial G_\alpha}{\partial e_\sigma} A^\alpha e_\beta C_{\eta\ \sigma}^{\ \beta} \dot{x}^k - G_\alpha\left(C_{\eta\ \varepsilon}^{\ \alpha} A^\varepsilon \dot{x}^k + C_{\beta\ \eta}^{\ \alpha} A^\beta \dot{x}^k\right) = 0$$

Because of the skew-symmetry (0.24) of the structure constants, this implies that

$$\frac{\partial G_\eta}{\partial e_\sigma}\left(F_\sigma + e_\beta C_{\mu\ \sigma}^{\ \beta} A^\mu\right)\dot{x}^k = -\frac{\partial G_\alpha}{\partial e_\sigma} A^\alpha e_\beta C_{\eta\ \sigma}^{\ \beta} \dot{x}^k$$

Furthermore, if k is chosen such that $\dot{x}^k \neq 0$, then we have

$$\frac{\partial G_\eta}{\partial e_\sigma}\left(F_\sigma + e_\beta C^{\beta}_{\mu\sigma} A^\mu\right) = -\frac{\partial G_\alpha}{\partial e_\sigma} A^\alpha e_\beta C^{\beta}_{\eta\sigma} \tag{5.20}$$

Alternatively, by contracting over k and p in (5.19) we obtain

$$\frac{\partial G_\eta}{\partial e_\sigma}\left(F_\sigma + e_\beta C^{\beta}_{\mu\sigma} A^\mu\right) n + \frac{\partial G_\alpha}{\partial e_\sigma} A^\alpha e_\beta C^{\beta}_{\eta\sigma} + (n-1) G_\alpha C^{\alpha}_{\eta\varepsilon} A^\varepsilon = 0$$

which upon substitution for the first term using (5.20) yields

$$(1-n)\frac{\partial G_\alpha}{\partial e_\sigma} A^\alpha e_\beta C^{\beta}_{\eta\sigma} + (n-1) G_\alpha C^{\alpha}_{\eta\varepsilon} A^\varepsilon = 0$$

Therefore, since $n > 1$, we have

$$\frac{\partial G_\alpha}{\partial e_\sigma} A^\alpha e_\beta C^{\beta}_{\eta\sigma} = G_\alpha C^{\alpha}_{\eta\beta} A^\beta \tag{5.21}$$

If this is differentiated with respect to u^μ_k there results

$$\frac{\partial G_\alpha}{\partial e_\sigma} e_\beta C^{\beta}_{\eta\sigma} \frac{\partial\left(G^\alpha_\nu \mathring{A}^\nu_h \dot{x}^h - \lambda^\alpha_\nu u^\nu_h \dot{x}^h\right)}{\partial u^\mu_k} = G_\alpha C^{\alpha}_{\eta\beta} \frac{\partial\left(G^\beta_\nu \mathring{A}^\nu_h \dot{x}^h - \lambda^\beta_\nu u^\nu_h \dot{x}^h\right)}{\partial u^\mu_k}$$

Since $\dot{x}^k \neq 0$, comparing coefficients of $\dot{x}^k$ shows that

$$\frac{\partial G_\alpha}{\partial e_\sigma} \lambda^\alpha_\mu e_\beta C^{\beta}_{\eta\sigma} = G_\alpha C^{\alpha}_{\eta\beta} \lambda^\beta_\mu$$

or, after multiplication by $\hat{\lambda}^\mu_\varepsilon\ \lambda^\eta_\gamma$

$$\frac{\partial G_\varepsilon}{\partial e_\sigma} e_\beta C^{\ \beta}_{\eta\ \sigma} \lambda^{\ \eta}_{\gamma} = G_\alpha C^{\ \alpha}_{\eta\ \varepsilon} \lambda^{\ \eta}_{\gamma}$$

According to (0.37), therefore

$$\frac{\partial G_\varepsilon}{\partial e_\sigma} \frac{\partial e_\sigma}{\partial u^\mu} = G_\alpha C^{\ \alpha}_{\eta\ \varepsilon} \lambda^{\ \eta}_{\mu}$$

so if we write $G_\alpha = G_\alpha(e_\sigma(u^\eta))$, it is observed that

$$\frac{\partial G_\varepsilon}{\partial u^\mu} = G_\alpha C^{\ \alpha}_{\eta\ \varepsilon} \lambda^{\ \eta}_{\mu}$$

Hence, we infer from (0.37) that *the quantities G_α constitute the components of a type (0,1) adjoint object* -

$$\boxed{G_\varepsilon = \hat{G}^{\ \alpha}_{\varepsilon} \overset{\circ}{G}_\alpha} \tag{5.22}$$

Substitution of (5.21) into (5.20) shows furthermore that

$$\frac{\partial G_\eta}{\partial e_\sigma}\left(F_\sigma + e_\beta C^{\ \beta}_{\mu\ \sigma} A^\mu_j \dot{x}^j\right) - G_\alpha C^{\ \alpha}_{\beta\ \eta} A^\beta_j \dot{x}^j = 0$$

and this together with (5.3) yields

$$\frac{\partial G_\eta}{\partial e_\sigma} f_\sigma - G_\alpha C^{\ \alpha}_{\beta\ \eta} A^\beta_j \dot{x}^j = 0 \tag{5.23}$$

Hence, *the absolute derivative of the adjoint type (0,1) quantity whose components are G_η vanishes*:

$$\boxed{\frac{DG_\eta}{D\tau} = 0} \tag{5.24}$$

Now substitute (5.15) into (2.34) to obtain (since $\gamma > 0$)

$$\left[\frac{\partial \mathring{G}_\alpha}{\partial \mathring{e}_\sigma} \mathring{A}^\alpha + \frac{\partial \mathring{C}}{\partial \mathring{e}_\sigma} \sqrt{\gamma} \right] \mathring{f}_\sigma = 0$$

Due to (5.23) and (1.3) this implies that

$$\mathring{G}_\eta \, C^{\eta}_{\beta\alpha} \mathring{A}^\beta \mathring{A}^\alpha + \frac{d\mathring{C}}{d\tau} \sqrt{\gamma} = 0$$

Because of the skew-symmetry of the structure constants the first term vanishes. Since also $\gamma \neq 0$, therefore

$$\frac{d\mathring{C}}{d\tau} = 0$$

and because of (5.17) we have

$$\boxed{\frac{dC}{d\tau} = 0} \tag{5.25}$$

According to (5.17) and (5.25), $C(e_\alpha(\tau, u^\sigma))$ does not depend upon u^σ or τ. Hence *C is a constant function of* $\mathring{e}_\alpha$.

We have thus shown that the assumptions of chapter 1 imply that L may be expressed in the form (5.15), where the functions G_η are C^3 scalars for which (5.22) and (5.24) hold and C is a nonzero constant function. Beginning at equation (5.16), it was also assumed that the signature of $[g_{ij}]$ is such that $\gamma > 0$; but the reader may observe that if $\gamma < 0$ the results (5.17), (5.22), (5.24) and (5.25) are still valid (by a completely analogous argument). These equations were obtained without appealing to the third invariance identity.

Lemma 2: *Let L be a C^3 scalar function of the form (1.7) defined on an admissible domain for which $g_{ij}\dot{x}^i\dot{x}^j > 0$, whose evaluation at $u = e$ is denoted by $\overset{\circ}{L}$. Assume that (1.9) holds, and that the fundamental integral (1.11) is parameter-invariant. Let $\overset{\circ}{E}_k$ denote the Euler-Lagrange vector (1.14), and let E_k be the vector obtained from $\overset{\circ}{E}_k$ by the replacement of $\overset{\circ}{e}_\alpha$, $\overset{\circ}{f}_\alpha$, $\overset{\circ}{A}{}^\alpha_j$, $\overset{\circ}{A}{}^\alpha_{j,k}$ with e_α, f_α, A^α_j, $A^\alpha_{j,k}$ respectively. Then in order that E_k be gauge-invariant it is necessary that the function L admit the representation*

$$\boxed{L = C\sqrt{g_{ij}\dot{x}^i\dot{x}^j} + G_\alpha A^\alpha_j \dot{x}^j}$$

for some nonzero constant function $C(\overset{\circ}{e}_\alpha)$ and C^3 scalar functions $G_\alpha = G_\alpha(e_\eta)$ which are such that

$$\boxed{G_\varepsilon = \hat{G}^\alpha_\varepsilon \overset{\circ}{G}_\alpha \quad ; \quad \frac{DG_\eta}{D\tau} = 0}$$

5.4 The Gauge-Invariant Equations of Motion

Let us apply the results of sections 5.2 and 5.3 to determine the equations of motion for a particle in a classical gauge field. We obtain the equations of motion from (5.15) and the conditions (5.17), (5.22), (5.24) and (5.25) in order to establish a converse to Lemma 2.

If equation (5.15) is evaluated at $u = e$, and we continue to assume that for all time-like vectors $\gamma > 0$, the Lagrangian is found to be

$$\mathring{L}^* = \mathring{C}\sqrt{g_{ij}\dot{x}^i\dot{x}^j} + \mathring{G}_\alpha \mathring{A}_j^\alpha \dot{x}^j$$

or, by (5.17)

$$\mathring{L}^* = C\sqrt{g_{ij}\dot{x}^i\dot{x}^j} + \mathring{G}_\alpha \mathring{A}_j^\alpha \dot{x}^j$$

Since (5.17) and (5.25) show that C is independent of τ and u^α, we use the notation

$$\mathring{L}^* = C\sqrt{g_{ij}(x^h)\dot{x}^i\dot{x}^j} + G_\alpha(\mathring{e}_\eta(\tau))\, \mathring{A}_j^\alpha(x^h)\, \dot{x}^j$$

and compute the derivatives required for the Euler-Lagrange vector:

$$\frac{\partial \mathring{L}^*}{\partial x^k} = C\,\frac{g_{ij,k}\dot{x}^i\dot{x}^j}{2\sqrt{\gamma}} + \mathring{G}_\alpha \mathring{A}_{j,k}^\alpha \dot{x}^j$$

$$\frac{\partial \mathring{L}^*}{\partial \dot{x}^k} = C\,\frac{g_{kq}\dot{x}^q}{\sqrt{\gamma}} + \mathring{G}_\alpha \mathring{A}_k^\alpha$$

and

$$\frac{d}{d\tau}\left(\frac{\partial \mathring{L}^*}{\partial \dot{x}^k}\right) = C\left(\frac{g_{kq,h}\dot{x}^q\dot{x}^h}{\sqrt{\gamma}} + \frac{g_{kq}\ddot{x}^q}{\sqrt{\gamma}} - \frac{g_{kq}\dot{x}^q}{2\gamma^{3/2}}\frac{d\gamma}{d\tau}\right) + \frac{d\mathring{G}_\alpha}{d\tau}\mathring{A}_k^\alpha + \mathring{G}_\alpha \mathring{A}_{k,h}^\alpha \dot{x}^h$$

Therefore the Euler-Lagrange vector is

$$\mathring{E}_k = \left[C\left(\frac{g_{kq,h}\dot{x}^q\dot{x}^h}{\sqrt{\gamma}} + \frac{g_{kq}\ddot{x}^q}{\sqrt{\gamma}} - \frac{g_{kq}\dot{x}^q}{2\gamma^{3/2}}\frac{d\gamma}{d\tau}\right) + \frac{d\mathring{G}_\alpha}{d\tau}\mathring{A}_k^\alpha + \mathring{G}_\alpha \mathring{A}_{k,h}^\alpha \dot{x}^h\right]$$
$$- \left[C\,\frac{g_{ij,k}\dot{x}^i\dot{x}^j}{2\sqrt{\gamma}} + \mathring{G}_\alpha \mathring{A}_{j,k}^\alpha \dot{x}^j\right]$$

or

$$\mathring{E}_k = C\left([g_{kq,h} + g_{hk,q} - g_{qh,k}]\frac{\dot{x}^q\dot{x}^h}{2\sqrt{\gamma}} + \frac{g_{kq}\ddot{x}^q}{\sqrt{\gamma}} - \frac{g_{kq}\dot{x}^q}{2\gamma^{3/2}}\frac{d\gamma}{d\tau}\right) + \frac{d\mathring{G}_\alpha}{d\tau}\mathring{A}_k^\alpha + \mathring{G}_\alpha\left[\mathring{A}_{k,h}^\alpha - \mathring{A}_{h,k}^\alpha\right]\dot{x}^h$$

$$= C\left(\frac{g_{kq}}{\sqrt{\gamma}}\frac{D\dot{x}^q}{D\tau} - \frac{g_{kq}\dot{x}^q}{2\gamma^{3/2}}\frac{d\gamma}{d\tau}\right) + \frac{d\mathring{G}_\alpha}{d\tau}\mathring{A}^\alpha_k + \mathring{G}_\alpha\left[\mathring{A}^\alpha_{k,h} - \mathring{A}^\alpha_{h,k}\right]\dot{x}^h$$

Equation (5.24) implies that we have

$$\mathring{E}_k = \frac{C g_{kq}}{\sqrt{\gamma}}\frac{D\dot{x}^q}{D\tau} + \mathring{G}_\eta C^\sigma_{\mu\alpha}\mathring{A}^\mu_j\mathring{A}^\alpha_k\dot{x}^j + \mathring{G}_\alpha\left[\mathring{A}^\alpha_{k,h} - \mathring{A}^\alpha_{h,k}\right]\dot{x}^h - C\frac{g_{kq}\dot{x}^q}{2\gamma^{3/2}}\frac{d\gamma}{d\tau}$$

Thus,

$$E_k = \frac{C g_{kq}}{\sqrt{\gamma}}\frac{D\dot{x}^q}{D\tau} + G_\eta C^\sigma_{\mu\alpha}A^\mu_j A^\alpha_k\dot{x}^j + G_\alpha\left[A^\alpha_{k,h} - A^\alpha_{h,k}\right]\dot{x}^h - \frac{C g_{kq}\dot{x}^q}{2\gamma^{3/2}}\frac{d\gamma}{d\tau}$$

so, by (0.51)

$$\boxed{E_k = \frac{C g_{kq}}{\sqrt{\gamma}}\frac{D\dot{x}^q}{D\tau} - G_\alpha F^\alpha_{h\,k}\dot{x}^h - \frac{C g_{kq}\dot{x}^q}{2\gamma^{3/2}}\frac{d\gamma}{d\tau}} \tag{5.26}$$

Due to (0.53), (1.1) and (5.17) each quantity shown is of adjoint type, and hence this expression is gauge-invariant. Our equations of motion (the Euler-Lagrange equations $\mathring{E}_k = 0$) are therefore equivalent to the equations $E_k = 0$, or

$$\boxed{g_{ik}\frac{D\dot{x}^i}{D\tau} = \sqrt{\gamma}\frac{G_\alpha}{C}F^\alpha_{j\,k}\dot{x}^j + \frac{g_{ik}\dot{x}^i}{2\gamma}\frac{d\gamma}{d\tau}} \tag{5.27}$$

We have shown that (5.15) together with the conditions (5.17), (5.22), (5.24) and (5.25) yields the gauge-invariant vector (5.26).[17] This establishes the converse of Lemma 2.

---

[17] If $\gamma < 0$ in equation (5.15), then a similar analysis shows that in equations (5.26) and (5.27) γ must be replaced by $|\gamma|$ and $\frac{D\dot{x}^i}{D\tau}$ must be replaced by $-\frac{D\dot{x}^i}{D\tau}$.

5.5 Summary of Results

Theorem 2: *Let L be a C^3 scalar function of the form*

$$L = L(e_\alpha, A_j^\alpha, g_{ij}, \dot{x}^j)$$

defined on an admissible domain for which $g_{ij}\dot{x}^i\dot{x}^j > 0$, whose evaluation at $u = e$ is denoted by $\overset{\circ}{L}$. Assume that (1.9) holds, and that the fundamental integral (1.11) is parameter-invariant. Let $\overset{\circ}{E}_k$ denote the Euler-Lagrange vector (1.14), and E_k the vector obtained from $\overset{\circ}{E}_k$ by replacing $\overset{\circ}{e}_\alpha, \overset{\circ}{f}_\alpha, \overset{\circ}{A}{}_j^\alpha, \overset{\circ}{A}{}_{j,k}^\alpha$ with $e_\alpha, f_\alpha, A_j^\alpha, A_{j,k}^\alpha$ respectively. Then E_k is gauge-invariant if and only if the function L admits the representation

$$\boxed{L = C\sqrt{g_{ij}\dot{x}^i\dot{x}^j} + G_\alpha A_j^\alpha \dot{x}^j}$$

for some nonzero constant[18] C and C^3 scalar functions $G_\alpha = G_\alpha(e_\eta)$ which are such that

$$\boxed{G_\varepsilon = \hat{G}_\varepsilon^\alpha \overset{\circ}{G}_\alpha \quad ; \quad \frac{DG_\eta}{D\tau} = 0}$$

In this case E_k is given by (5.26), and the Euler-Lagrange equations are equivalent to the gauge-invariant equations (5.27).

Note that if τ is chosen to be proper time, equations (5.26) and (5.27) become remarkably simple, i.e.

[18] The value of C may depend upon the coupling parameters, i.e. $C = C(\overset{\circ}{e}_\alpha)$, but is a constant for any given coupling parameters.

$$\boxed{E_k = C\, g_{ik}\frac{D\dot{x}^i}{D\tau} - G_\alpha F_{j\,k}^{\alpha}\dot{x}^j} \qquad (5.28)$$

and

$$\boxed{g_{ik}\frac{D\dot{x}^i}{D\tau} = \frac{G_\alpha}{C} F_{j\,k}^{\alpha}\dot{x}^j} \qquad (5.29)$$

The vanishing of the absolute derivative (5.24) is in agreement with certain approaches to classical gauge field theory. For instance Rund [13], p. 113, obtains a similar result on flat (Minkowski) space-time.[19] If G is the unitary group $U(1)$, while M is Minkowski space-time, $C = mc$ and G_1 is the charge of the particle, then the Lagrangian in Theorem 2 is the usual Lagrangian of electromagnetic theory. Equations (5.24) in this case, state that the electric charge of a particle is a constant. Wong [1], pp. 691-693, obtains a relationship which is formally similar to (5.24) in the classical limit of a quantum mechanical result for certain entities associated with the generators of $SU(2)$.

Finally, because of the gauge-invariance of the equations $E_k = 0$, Corollary 2 implies that equations (5.27) together with arbitrary initial values (3.12) in the domain of L have a unique solution on an interval $\tau \in (\tau_o - \varepsilon, \tau_o + \varepsilon)$ for some real number $\varepsilon > 0$.

These strong results were obtained as a consequence of the use of functional equations to determine how the arguments of the Lagrangian must combine (chapter 2). Without such an understanding of the form of the Lagrangian, our later reasoning would not have been possible.

[19] However, in that paper the standard form of the free-field Lagrangian is adopted ab initio, and it is also assumed that the full Lagrangian is the sum of the free-field Lagrangian and an interaction Lagrangian that is linear in the velocity components.

APPENDIX: A USEFUL FACT

The purpose of this appendix is to establish the following fact (with use of the summation convention suspended).

Fact: *Let f be a differentiable function of N independent real variables $y^1, y^2, \ldots, y^N$, defined on an open region A of R^N. Then, f satisfies the system of partial differential equations*

$$\boxed{\frac{\partial f}{\partial y^I} - \frac{\partial f}{\partial y^J} = 0} \tag{A.1}$$

(in which I,J represent arbitrary integers between 1 and N) on A if and only if there exists a differentiable function F of a single variable, such that on A

$$\boxed{f = F\left(\sum_{I=1}^{N} y^I\right)} \tag{A.2}$$

The proof of this follows from the lemma below.

Lemma: *Let f be a differentiable function of N independent real variables $y^1, y^2, \ldots, y^N$ with $N > 1$, defined on an open region A of R^N. Then f satisfies the system of partial differential equations*

$$\frac{\partial f}{\partial y^I} - \frac{\partial f}{\partial y^J} = 0$$

(in which I,J represent arbitrary integers between 1 and N) on A if and only if there exists

a differentiable function h of the N-1 quantities

$$z^1 = y^1 \ , \ \dots \ , \ z^{N-2} = y^{N-2} \ , \ z^{N-1} = y^{N-1} + y^N \ , \ z^N = y^{N-1} - y^N$$

such that on A,

$$f \ = \ h\,(z^1 \ , \ \dots \ , \ z^{N-1}) \quad \textit{and} \quad \frac{\partial h}{\partial z^I} - \frac{\partial h}{\partial z^J} = 0$$

where I,J = 1,...,N-1.

Proof of the lemma:

Consider the change of variables

$$z^1 = y^1 \ , \ \dots \ , \ z^{N-2} = y^{N-2} \ , \ z^{N-1} = y^{N-1} + y^N \ , \ z^N = y^{N-1} - y^N \qquad \text{(A.3)}$$

Because this change of variables is invertible there exists a unique differentiable function g such that

$$f(y^1 \ , \ \dots \ , \ y^N) \ = \ g\,(z^1 \ , \ \dots \ , \ z^N)$$

Let us substitute g into the last of the equations (A.1) (in which $I = N$ and $J = N$-1). We get

$$\frac{\partial f}{\partial y^N} - \frac{\partial f}{\partial y^{N-1}} = 0$$

The left-hand side of this equation is

$$\frac{\partial f}{\partial y^N} - \frac{\partial f}{\partial y^{N-1}} = \left(\frac{\partial g}{\partial z^{N-1}}\frac{\partial z^{N-1}}{\partial y^N} + \frac{\partial g}{\partial z^N}\frac{\partial z^N}{\partial y^N}\right) - \left(\frac{\partial g}{\partial z^{N-1}}\frac{\partial z^{N-1}}{\partial y^{N-1}} + \frac{\partial g}{\partial z^N}\frac{\partial z^N}{\partial y^{N-1}}\right)$$

$$= \left(\frac{\partial g}{\partial z^{N-1}} - \frac{\partial g}{\partial z^N}\right) - \left(\frac{\partial g}{\partial z^{N-1}} + \frac{\partial g}{\partial z^N}\right)$$

$$= -2\frac{\partial g}{\partial z^N}$$

so we have in terms of g, that

$$\frac{\partial g}{\partial z^N} = 0$$

This holds if and only if there exists a differentiable function h of N-1 variables such that

$$f(y^1, \dots, y^N) = g(z^1, \dots, z^N) = h(z^1, \dots, z^{N-1}) \tag{A.4}$$

The remaining equations (A.1) may now be computed in terms of h as follows. Differentiation of (A.4) reveals that (due to (A.3))

$$\frac{\partial f}{\partial y^I} = \sum_{J=1}^{N-1} \frac{\partial h}{\partial z^J}\frac{\partial z^J}{\partial y^I}$$

$$= \begin{cases} \dfrac{\partial h}{\partial z^I} & (\text{if } I < N-1) \\ \dfrac{\partial h}{\partial z^{N-1}}\dfrac{\partial (y^{N-1} + y^N)}{\partial y^I} & (\text{if } I = N-1, \text{ or } I = N) \end{cases}$$

or

$$\frac{\partial f}{\partial y^I} = \begin{cases} \dfrac{\partial h}{\partial z^I} & \text{(if } I < N\text{-}1) \\[2ex] \dfrac{\partial h}{\partial z^{N-1}} & \text{(if } I = N\text{-}1 \text{ or } I = N) \end{cases}$$

so the equations (A.1) are equivalent to the system of equations

$$\frac{\partial h}{\partial z^I} - \frac{\partial h}{\partial z^J} = 0 \tag{A.5}$$

where I,J now vary only over $1,\ldots,N\text{-}1$. Thus, satisfaction of the system (A.1) is equivalent to the requirement that there exist a differentiable function h which depends only on the N-1 variables $z^1,\ldots,z^{N-1}$ shown in (A.3), and which satisfies the equation (A.4) and the system of differential equations (A.5).

Proof of the fact:

If $N = 1$, there is nothing to be proved. If $N > 1$, we apply the lemma to the function f. The lemma shows that f is equal to a function h (as in (A.4)) of N-1 independent variables, which must satisfy the system (A.5) of differential equations if $N>2$. If h is not a function of a single variable, then because (A.5) is analogous to (A.1) the lemma applies to h. Hence it is deduced that there exists a differentiable function $\tilde{h}$ of the N-2 variables:

$$a^1 = z^1 ,\ldots, a^{N-3} = z^{N-3} , a^{N-2} = z^{N-2} + z^{N-1}$$

that is, of

$$a^1 = y^1 ,\ldots, a^{N-3} = y^{N-3} , a^{N-2} = y^{N-2} + y^{N-1} + y^N$$

such that $h = \tilde{h}$. If $\tilde{h}$ is not a function of one variable, it satisfies a system of differential equations (in terms of N-2 variables) analogous to (A.1) and the lemma may be applied again. This iteration may be repeated until a function of the single variable

$$\sum_{I=1}^{N} y^I$$

is obtained. Since the function obtained by this process is equal to f, the fact is established.

REFERENCES

1. Wong, S. K., Field and Particle Equations for the Classical Yang-Mills Field and Particles with Isotopic Spin, *Nuovo Cimento* **65** (A), 689-694 (1970).

2. Lovelock, D. and Rund, H., "Tensors, Differential Forms, and Variational Principles." Wiley (Interscience), New York, 1975; Dover reprint, New York 1989.

3. Drechsler, W. and Mayer, M. E., "Fiber Bundle Techniques in Gauge Theories" (Lecture Notes in Physics). Springer-Verlag, Berlin, 1977.

4. Rund, H., Differential-Geometric and Variational Background of Classical Gauge Field Theories, *Aequationes Mathematicae* **24**, 121-174 (1982).

5. Pontryagin, L. S., "Topological Groups" (Second Ed.). Gordon and Breach, New York, 1966.

6. Sagle, A. and Walde, R., "Introduction to Lie Groups and Lie Algebras." Academic Press, New York, 1973.

7. Greub, W., Halperin, S. and Vanstone, R., "Connections, Curvature and Cohomology," Volume 2. Academic Press, New York, 1973.

8. Rund, H., Adjoint Connections on Group Manifolds and Gauge Transformations, *The Mathematical Heritage of C.F. Gauss: Geometry, Topology, and Number Theory*, ed. G.M. Rassias, Teubner, Leipzig (to be published).

9. Logan, J. D., "Invariant Variational Principles" (Mathematics in Science and Engineering, Vol. 138). Academic Press, New York, 1973.

10. Aczel, J., "Lectures on Functional Equations and Their Applications." Academic Press, New York, 1966.

11. Rund, H., "The Hamilton Jacobi Theory in the Calculus of Variations." Van Nostrand, London and New York, 1966; Krieger reprint, Huntington, N.Y., 1973.

12. Pontryagin, L. S., "Ordinary Differential Equations." Addison-Wesley, Reading, Mass., 1962.

13. Rund, H., Invariance Identities Associated with Finite Gauge Transformations and the Uniqueness of the Equations of Motion of a Particle in a Classical Gauge Field, *Foundations of Physics* **13** (1), 93-114 (Jan. 1983).