International Journal of Theoretical Physics, Vol. 24, No. 5, 1985

Perturbed Characteristic Functions, II Second-Order Perturbation

H. A. Buchdahl[1,2]

Received October 27, 1984

In a previous paper I considered the first-order perturbation $V^{(1)}$ of characteristic functions, perturbation of the world characteristic serving as illustration. In various contexts, knowledge of $V^{(1)}$ is, however, insufficient already in principle to afford even an approximate solution of some physical problem. I therefore go on now to an investigation of the second-order perturbation $V^{(2)}$. To find it one not unexpectedly needs to know only the *first*-order perturbation of the extremal. The general theory is illustrated by the detailed calculations giving the world characteristic of a particular class of metrics, correct to within third-order terms.

1. INTRODUCTION

Let $L(\dot{q}, q, u)$ be a function of a set of functions $q(u) =: \{q^k(u);\ k = 1, 2, \ldots, n\}$, their first derivatives $\dot{q}(u) =: \{\dot{q}^k(u);\ k = 1, 2, \ldots, n\}$ and of the independent variable u explicitly. If t and t' are a pair of fixed, but arbitrarily selected, terminal values of u, the corresponding values of the $q^k(u)$ will be $x^k := q^k(t)$ and $x^{k'} := q^k(t')$. $\{x^1, x^2, \ldots, x^n, t\}$ and $\{x^{1'}, x^{2'}, \ldots, x^{n'}, t'\}$ may be thought of as the coordinates, in a representative space R_{n+1}, of the endpoints A, A' of the curve $\{q(u), t \le u \le t'\}$ joining A and A'. Henceforth let this curve in fact be the extremal E, assumed unique, which passes through A and A', i.e., that curve which extremizes the functional

$$\tilde{V} := \int_t^{t'} L(\dot{q}, q, u)\, du \tag{1}$$

When $\tilde{V}$ is evaluated along E it becomes a function V of $2n+2$ variables, namely, of the coordinates of A and A'. This is the *characteristic function*

[1]Institut für theoretische Physik, Universität Bern, CH-3012, Bern, Switzerland.
[2]Permanent address: Department of Physics and Theoretical Physics, Australian National University, Canberra 6201, Australia.

0020-7748/85/0500-0457$04.50/0

which belongs to the "Lagrangian" L:

$$V(x', t', x, t) := \int_E L\, du \tag{2}$$

where $x' := \{x^{k'}\}$, $x := \{x^k\}$. It usually goes under different names in various branches of physics: the point characteristic in geometrical optics, the principal function in analytical mechanics, while in general relativity theory I have called it (Buchdahl, 1979) the world characteristic because $\Omega := \frac{1}{2}|V^2|$ is Synge's world function (Synge, 1960).

The characteristic function which belongs to any particular problem constitutes its solution. It is closely connected with the Hamilton–Jacobi theory: it satisfies a Hamilton–Jacobi equation both at A and at A'; and, unlike Ω, it is in fact the generator of the canonical transformation which transforms the canonical variables from their initial to their final values. While V thus evidently occupies a central position within the various theories, its practical usefulness is severely circumscribed because in all but a handful of highly specialized or contrived cases (e.g., Buchdahl, 1970, 1973, 1975) its explicit evaluation is not feasible. However, it is not unusual to encounter situations in which the Lagrangian L differs "only little" from a Lagrangian $L^{(0)}$ whose associated characteristic function $V^{(0)}$ is known explicitly. Thus, let it be granted that there is a parameter ε, sufficiently small in absolute value, such that

$$L = L^{(0)} + \varepsilon L^{(1)} + \varepsilon^2 L^{(2)} + \cdots \tag{3}$$

the series on the right being convergent. Then, correspondingly,

$$V = V^{(0)} + \varepsilon V^{(1)} + \varepsilon^2 V^{(2)} + \cdots \tag{4}$$

cases in which (4) does not obtain being here excluded from consideration. In short, one has a perturbation problem, namely, that of devising a method for finding the "perturbations" $V^{(1)}$, $V^{(2)}, \ldots$, in turn.

In (Buchdahl, 1979) only $V^{(1)}$ was considered, mainly in the context of gravitational fields due to weak sources. There one has a rather simple state of affairs for two reasons: (i) quite generally, $V^{(1)}$ is the integral of $L^{(1)}$ along the *unperturbed* extremal and (ii) since $V^{(0)}$ corresponds to flat space it is known immediately, i.e.,

$$V^{(0)} = (\eta_{ij} h^i h^j)^{1/2} \tag{5}$$

where $\eta_{ij} = \mathrm{diag}(1, 1, 1, -1)$ and $h^i := x^{i'} - x^i$. In these circumstances some quite striking general (if formal) results may be obtained.

Now, under various circumstances knowledge of $V^{(1)}$ alone is, in the nature of things, inadequate for the (approximate) solution of some physical problem. For instance, one cannot, even in principle, describe achromatization in optics without taking at least also $V^{(2)}$ into account; likewise, in a region in which the Schwarzchild field may be looked upon as a perturbation of flat space, the calculation of the precession of planetary perihelia cannot go through without a knowledge of $V^{(2)}$. This paper therefore addresses itself to the general problem of finding $V^{(2)}$. Not unexpectedly it turns out in Section 2 that to do so one needs to know only the *first*-order perturbation of the extremal. Section 3 serves as a reminder that the first-order perturbation of the extremal is itself contained in $V^{(0)}+\varepsilon V^{(1)}$, already known. Section 4 deals with the world characteristic in particular, special attention being paid to the case in which the unperturbed space is flat. By way of illustrating the generic results obtained, the actual calculations are carried out in detail for a particular class of metrics in Section 5.

2. GENERIC EXPRESSIONS FOR $V^{(2)}$

Essentially one is concerned with the difference $\Delta V := V - V^{(0)}$ between the two characteristic functions

$$V(x', t', x, t) = \int_E L\, du \tag{6}$$

and

$$V^{(0)}(x', t', x, t) = \int_{E_0} L^{(0)}\, du \tag{7}$$

where the integration are, respectively, along the extremals E and E_0 which belong to L and $L^{(0)}$, both E and E_0 passing through A and A'. If $\{\bar{q}^1, \ldots, \bar{q}^n, u\}$ are coordinates in R_{n+1}, the equations of E are

$$\bar{q}^k = q^k(u) \tag{8}$$

and those of E_0 are

$$\bar{q}^k = q^k_{(0)}(u) \tag{9}$$

say. Consistently with the assumption already made that V can be written as a series in ascending powers of ε, one requires that (8) can be written

$$\begin{aligned}\bar{q}^k &= q^k_{(0)}(u) + \varepsilon q^k_{(1)}(u) + \varepsilon^2 q^k_{(2)}(u) + \cdots \\ &=: q^k_{(0)}(u) + \chi^k(u)\end{aligned} \tag{10}$$

say, Then

$$\Delta V = \int \{L(\dot{q}_{(0)} + \dot{\chi}, q_{(0)} + \chi, u) - L^{(0)}(\dot{q}_{(0)}, q_{(0)}, u)\}\, du \tag{11}$$

The notation may be further simplified by agreeing to the convention that when (and only when) arguments of Lagrangians, of their derivatives, or, later, of metric tensors are omitted they are to be taken as referring to unperturbed extremals, e.g., L stands for $L(\dot{q}_{(0)}, q_{(0)}, u)$. If subscripts denote derivatives, (11) now becomes

$$\Delta V = \int \{L^* + (L_{\dot{q}}\dot{\chi} + L_q\chi) + (\tfrac{1}{2}L_{\dot{q}\dot{q}}\dot{\chi}^2 + L_{\dot{q}q}\dot{\chi}\chi + \tfrac{1}{2}L_{qq}\chi^2) + \cdots\}\, du \quad (12)$$

when $L^* = L - L^{(0)}$. This equation has been written as if $n = 1$. Its interpretation when $n > 1$ is almost self-evident; e.g., $L_{\dot{q}\dot{q}}\dot{\chi}^2$ stands for $\sum_{i,j} \dot{\chi}^i\dot{\chi}^j \partial^2 L/\partial\dot{q}^i\partial\dot{q}^j$.

Now recall that the $q^k(u)$ describe an extremal. The extremality implies in the usual way that

$$\int_t^{t'} \{\dot{\eta}L_{\dot{q}}(\dot{q}, q, u) + \eta L_q(\dot{q}, q, u)\}\, du = 0 \quad (13)$$

if only the otherwise arbitrary functions η $(:= \{\eta^k(u)\})$ vanish at A and A'. The choice $\eta = \chi$ is therefore permissible. Using also (10), (13) then becomes

$$\int_t^{t'} \{\dot{\chi}[L_{\dot{q}} + (\dot{\chi}L_{\dot{q}\dot{q}} + \chi L_{\dot{q}q}) + \cdots] + \chi[L_q + (\dot{\chi}L_{\dot{q}q} + \chi L_{qq}) + \cdots]\}\, du = 0 \quad (14)$$

Subtraction from (12) leads to the relation

$$\Delta V = \int_t^{t'} \{L^* - (\tfrac{1}{2}\dot{\chi}^2 L_{\dot{q}\dot{q}} + \dot{\chi}\chi L_{\dot{q}q} + \tfrac{1}{2}\chi^2 L_{qq}) + \cdots\}\, du \quad (14a)$$

i.e.,

$$\Delta V = \int_t^{t'} \{\varepsilon L^{(1)} + \varepsilon^2[L^{(2)} - (\tfrac{1}{2}\dot{\chi}^2 L^{(0)}_{\dot{q}\dot{q}} + \dot{\chi}\chi L^{(0)}_{\dot{q}q} + \tfrac{1}{2}\chi^2 L^{(0)}_{qq})]\}\, du + O(\varepsilon^3)$$

Thus

$$V^{(1)} = \int_t^{t'} L^{(1)}\, du \quad (15)$$

as in Buchdahl (1979), while the required relation for $V^{(2)}$ is

$$V^{(2)} = \int_t^{t'} \{L^{(2)} - (\tfrac{1}{2}L^{(0)}_{\dot{q}\dot{q}}\dot{q}^2_{(1)} + L^{(0)}_{\dot{q}q}\dot{q}_{(1)}q_{(1)} + \tfrac{1}{2}L^{(0)}_{qq}q^2_{(1)})\}\, du \quad (16)$$

Clearly one does not need to know the functions $q_{(2)}(u)$ here. (16) may be cast into an interesting alternative form as follows. Write (14) as

$$\int_t^{t'} \{(\dot\chi L_{\dot q}^{(0)} + \chi L_q^{(0)}) + (\dot\chi L_{\dot q}^* + \chi L_q^*) + (\dot\chi^2 L_{\dot q\dot q} + 2\dot\chi\chi L_{\dot q q} + \chi^2 L_{qq})\}\, du = O(\varepsilon^3)$$

The first two terms do not contribute to the integral on account of the extremality of E_0 [cf. equation (13)]. One is left with

$$\int_t^{t'} (\dot q_{(1)} L_{\dot q}^{(1)} + q_{(1)} L_q^{(1)})\, du = -\int_t^{t'} (\dot q_{(1)}^2 L_{\dot q\dot q}^{(0)} + 2\dot q_{(1)} q_{(1)} L_{\dot q q}^{(0)} + q_{(1)}^2 L_{qq}^{(0)})\, du$$

to within terms of order ε^3. Therefore

$$V^{(2)} = \int_t^{t'} \{L^{(2)} + \tfrac{1}{2}(\dot q_{(1)} L_{\dot q}^{(1)} + q_{(1)} L_q^{(1)})\}\, du \tag{17}$$

3. REMARK ON $q_1(u)$

Equations (15) and (16) constitute a self-contained pair of equations in the sense that the functions $q_{(1)}$ required in (16) are themselves directly derivable from $V^{(0)} + \varepsilon V^{(1)}$ ($=: \hat V$, say). To this end one need only let $\hat V(q, u, x, t)$ assume its function as the generator of the canonical transformation connecting A with current points $\bar A$ on E. In other words, one solves (correctly to the first order in ε) the n equations

$$\partial_i\{\hat V(q, u, x, t) - \hat V(x', t', x, t)\} = 0 \tag{18}$$

($\partial_i := \partial/\partial x^i$) for the q^i as functions of u and, of course, x', t', x, t.

4. THE CASE OF THE WORLD CHARACTERISTIC

The world characteristic V belongs to the Lagrangian

$$L = (\dot q^i \dot q_i)^{1/2} \tag{19}$$

where, of course, $\dot q_i = g_{ij}\dot q^j$, g_{ij} being the metric tensor of a four-dimensional Riemann space of signature 2. (Actually the signature and dimensionality are here largely irrelevant.) In the present context g_{ij} is assumed to be a real-analytic function of a parameter ε in a neighborhood of $\varepsilon = 0$, i.e., when $|\varepsilon|$ is sufficiently small g_{ij} can be represented as a power series

$$g_{ij} = g_{ij}^{(0)} + \varepsilon g_{ij}^{(1)} + \varepsilon^2 g_{ij}^{(2)} + \cdots \tag{20}$$

the first term on the right being the "base metric." [Note that indices in parentheses, characterizing the order (of perturbation) will, as a matter of convenience, be indiscriminately moved up or down; e.g., the inverse of $g^{(0)}_{ij}$ is $g^{ij}_{(0)}$.] Then

$$L^{(0)}(\dot{q}, q, u) = [g^{(0)}_{ij}(q)\dot{q}^i\dot{q}^j]^{1/2}$$
$$L^{(1)}(\dot{q}, q, u) = \tfrac{1}{2}g^{(1)}_{ij}(q)\dot{q}^i\dot{q}^j / L^{(0)}(\dot{q}, q, u) \tag{21}$$

and so on.

The important case of the flat base metric, $g^{(0)}_{ij} = \eta_{ij}$, say, represents a particularly simple state of affairs, for them the $q^i_{(0)}$ are linear functions of u. However this parameter was chosen originally, one can arrange it to have the value 0 at A and 1 at A'. No formal changes occur in the equations already derived on account of the homogeneity of L, except in as far as the limits of integration are now 0, 1 in place of t, t'. In short,

$$q^i_{(0)}(u) = h^i u + x^i \tag{22}$$

If indices be now lowered with η_{ij} instead of g_{ij} one therefore has $L^{(0)} = (h_i h^i)^{1/2} =: l$, which, being constant, trivially implies (5). At the same time, of the derivatives of $L^{(0)}$ which appear in (16) only the first does not vanish identically and it takes the form

$$L^{(0)}_{\dot{q}^i\dot{q}^j} = l^{-1}(\eta_{ij} - \bar{h}_i\bar{h}_j) \tag{23}$$

with $\bar{h}^i = l^{-1}h^i$. (16) now reduces to

$$V^{(2)} = \int_0^1 \{L^{(2)} - \tfrac{1}{2}l^{-1}[\dot{q}^{(1)}_i\dot{q}^i_{(1)} - (\bar{h}_i\dot{q}^i_{(1)})^2]\}\, du \tag{24}$$

whereas there are no analogous general consequences for (17). On the other hand, in the more specialized case in which the perturbed V_4 is conformally flat, i.e.,

$$g^{(r)}_{ij} = \psi^{(r)}\eta_{ij} \qquad (r = 1, 2, \ldots) \tag{25}$$

where the $\psi^{(r)}$ are scalars, it is only (17) that can be usefully rewritten. Thus now,

$$L^{(1)} = \tfrac{1}{2}l\psi^{(1)}, \qquad L^{(2)} = \tfrac{1}{2}l[\psi^{(2)} - \tfrac{1}{4}(\psi^{(1)})^2] \tag{26}$$

and, with $\partial_i := \partial/\partial q^i$ now,

$$L^{(1)}_{\dot{q}^i} = \tfrac{1}{2}\psi^{(1)}h_i, \qquad L^{(1)}_{q^i} = \tfrac{1}{2}l\partial_i\psi^{(1)} \tag{27}$$

After an integration by parts, to remove the derivative of $q_{(1)}$, (17) finally becomes

$$V^{(2)} = \tfrac{1}{2}l\int_0^1 \{[\psi^{(2)} - \tfrac{1}{4}(\psi^{(1)})^2] + \tfrac{1}{2}(\eta_{ij} - \bar{h}_i\bar{h}_j)q^i_{(1)}\partial^j\psi^{(1)}\}\, du \tag{28}$$

The projection tensor $\eta_{ij} - \bar{h}_i \bar{h}_j$ which appeared in (23) has turned up again here.

5. EXPLICIT EXAMPLE

The generic results obtained above are now to be applied to a tangible example. To keep the work as simple as possible I take the V_4 to be a "conformally perturbed flat space" the metric of which is, specifically,

$$ds^2 = (1 + \varepsilon q^1)^{-2} \eta_{ij} \, dq^i \, dq^j \tag{29}$$

(It is to be taken for granted that only points of the V_4 which have $|\varepsilon q^1| \ll 1$ are contemplated.) The perturbation of the metric is

$$g_{ij} - \eta_{ij} = [(1 + \varepsilon q^1)^{-2} - 1] \eta_{ij}$$

whence

$$\psi^{(r)} = (-1)^r (r+1)(q^1)^r \tag{30}$$

Then, from (15), (22), (26) and (30)

$$V^{(1)} = \tfrac{1}{2} l \int_0^1 \psi^{(1)} \, du = -l \int_0^1 (h^1 u + x^1) \, du$$

i.e.,

$$V^{(1)} = -\tfrac{1}{2} l (x^{1'} + x^1) \tag{31}$$

Thus

$$\hat{V} = l[1 - \tfrac{1}{2}\varepsilon(x^{1'} + x^1)] \tag{32}$$

Now let the indices a, b take only the values 2, 3, 4. Then

$$\begin{aligned} -\partial_1 \hat{V} &= l^{-1}\{(x^{1'} - x^1) + \tfrac{1}{2}\varepsilon[l^2 - (x^{1'})^2 + (x^1)^2]\} \\ -\partial_a \hat{V} &= l^{-1}(x^{a'} - x^a)[1 - \tfrac{1}{2}\varepsilon(x^{1'} + x^1)]. \end{aligned} \tag{33a–d}$$

The equations for the unperturbed geodesic are then given to the required order by (18), i.e., each of the four expressions on the right of (33a–d) is to be equated to the corresponding expression obtained from it by replacing x^j by q^j. In this process l becomes $\bar{l}$, say. Because of (22)

$$\bar{l} = lu + O(\varepsilon) \tag{34}$$

Of the four equations one must be redundant, bearing in mind that V satisfies the Hamilton-Jacobi equation. Indeed, upon inserting the expressions for $q^i - x^i$ in $\eta_{ij}(q^i - x^i)(q^j - x^j)$ the latter becomes $\bar{l}^2 + O(\varepsilon^2)$, i.e., to the required order one has an identity. One is therefore free to stipulate that

$$q^1(u) = h^1 u + x^1 \tag{35}$$

exactly. From (33a) one then infers directly that

$$\bar{l} = lu\{1 - \tfrac{1}{2}\varepsilon(d^2/h^1)(1-u) + O(\varepsilon^2)\} \tag{36}$$

with

$$d^2 := l^2 - (h^1)^2 = \eta_{ab} h^a h^b \tag{37}$$

The remaining equations then assume the simple form

$$q^a(u) = (h^a u + x^a) - \tfrac{1}{2}\varepsilon l^2 (h^a/h^1) u(1-u) + O(\varepsilon^2) \tag{38}$$

Thus

$$q^i_{(0)} = h^i u + x^i \tag{39}$$

and

$$q^1_{(1)}(u) = 0, \qquad q^a_{(1)}(u) = -\tfrac{1}{2} l^2 (h^a/h^1) u(1-u) \tag{40}$$

From these, by inspection,

$$\dot{q}^i_{(1)} \dot{q}^{(1)}_i = \tfrac{1}{4} l^4 (d/h^1)^2 u^2 (1-u)^2$$

and

$$\bar{h}_i \dot{q}^i_{(1)} = l(d^2/h^1)(u - \tfrac{1}{2})$$

Further, from (26) and (30)

$$L^{(2)} = l(q^1)^2$$

$V^{(2)}$ may now be calculated from (24):

$$V^{(2)} = l \int_0^1 \{(h^1 u + x^1)^2 - \tfrac{1}{2} d^2 (u - \tfrac{1}{2})^2\}\, du$$

The integration is trivial and gives

$$V^{(2)} = l\{\tfrac{1}{3}[(x^{1'})^2 + x^{1'}x^1 + (x^1)^2] - \tfrac{1}{24} d^2\} \tag{41}$$

The same result is obtained, perhaps even more easily, from (28). At any rate, the final result is

$$V = l\{1 - \tfrac{1}{2}\varepsilon(x^{1'} + x^1) + \tfrac{1}{3}\varepsilon^2[(x^{1'})^2 + x^{1'}x^1 + (x^1)^2 - \tfrac{1}{8}d^2] + O(\varepsilon^3)\} \tag{42}$$

One has a secure check upon this result since the characteristic function which belongs to the metric (29) can be found in closed form by standard methods. Thus

$$V = 2\varepsilon^{-1}\,\text{arsinh}[\tfrac{1}{2}\varepsilon l(1+\varepsilon x^{1'})^{-1/2}(1+\varepsilon x^{1})^{-1/2}] \tag{43}$$

It is easily confirmed that this is in harmony with (42).

One might gain the impression that the ease with which (42) could be obtained was somehow connected with the fact that V could in this case be obtained explicitly in terms of known functions. To dispel this notion it will suffice to consider the more general metric

$$ds^2 = (1+\varepsilon q^1)^{-2N}\eta_{ij}\,dx^i\,dx^j \tag{44}$$

where N is any real number. In general one will not be able to find V in closed form, whereas the changes required in equations (30)–(41) to accomodate this more general case are largely trivial. Wherever only first-order terms are involved one simply replaces ε by $N\varepsilon$, while $L^{(2)}$ takes an additional factor $\frac{1}{2}N(N+1)$. Then, in place of (42),

$$\begin{aligned} V = l(1 &- \tfrac{1}{2}N(x^{1'}+x^1)\varepsilon + \tfrac{1}{3}N\{\tfrac{1}{2}(N+1)[(x^{1'})^2 + x^{1'}x^1 + (x^2)^2] \\ &- \tfrac{1}{8}Nd^2\}\varepsilon^2 + O(\varepsilon^3)) \end{aligned} \tag{45}$$

6. CONCLUDING REMARK

It has already been noted that to find $V^{(1)}$ and $V^{(2)}$ one only needs to know $q(u)$ to the zeroth and first orders, respectively. This state of affairs generalizes to all orders as a consequence of the absence from the integral in (14a) of terms linear in χ and $\dot{\chi}$. In other words, to determine $V^{(n)}$ one needs to know only $q_{(r)}(u)$, $r = 1, 2, \ldots, n-1$.

REFERENCES

Buchdahl, H. A. (1970). *Optica Acta,* **17**, 707–713.
Buchdahl, H. A. (1973). *Gen. Rel. Grav.,* **3**, 35–41.
Buchdahl, H. A. (1975). *Optik,* **43**, 259–274.
Buchdahl, H. A. (1979). *Australian J. Phys.* **32**, 405–410.
Synge, J. L. (1960). *Relativity: The General Theory,* Chap. II., North-Holland Amsterdam.