X-641-69-515

PREPRINT

NASA TM X- 63 753

ON KRUSKAL'S PERTURBATION METHOD

DAVID P. STERN

DECEMBER 1969

GSFC

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

On Kruskal's Perturbation Method

David P. Stern
Laboratory for Theoretical Studies
Goddard Space Flight Center
Greenbelt, Maryland 20771

__Abstract__

A general method for eliminating the angle variable from the equations of a perturbed periodic motion and for deriving an "adiabatic invariant" J has been given by Kruskal, and for a special class of Hamiltonian systems, McNamara and Whiteman have shown (to order ^{2}) that J is related to a set of invariants I obtained from the expansion of Poisson bracket relations. In this work, an order-by-order algorithm for Kruskal's method is introduced and a new set of invariants Z_1 is obtained. It is shown that these invariants bear a close relation to those obtained from the Poisson bracket expansion, and in the special case investigated by McNamara and Whiteman, the relation between I and Z_1 may be brought to the same form as the relation between I and J derived by those authors. Finally, the relationship between Z_1 and J is examined and arguments are presented that in certain cases the two are equal to all orders.

INTRODUCTION

Let a perturbed periodic mechanical system be given, described by n canonical variables collectively represented by the vector $\underline{y}$ and by a Hamiltonian H dependent on a small parameter , of the form

$$H = y_1 + \sum_{k=1}^{\infty} \varepsilon^k H^{(k)}(\underline{y}) \qquad (1)$$

(here and in what follows, superscripts in parentheses denote order in ε). The vector $\underline{y}$ may be viewed as the solution of the Hamilton-Jacobi equation for the unperturbed motion, yielding an action variable y_1 , its conjugate angle variable y_n and a set of other variables y_i which are constants of the unperturbed motion.

A perturbation expansion may now be employed to eliminate y_1 and y_n from the equations of motion to any desired order in ε . One such technique has been devised by Kruskal[1] and will be described in more detail further on (this method is also applicable to non-canonical systems). By Kruskal's approach, a near-identity transformation from the variables $\underline{y}$ to new "nice" variables $\underline{z}$ is performed, so that of the n first-order differential equations describing the evolution of $\underline{z}$, (n-1) may be separated and solved independently as an autonomous set. Furthermore, using these "nice variables", an "adiabatic invariant" J ,which is a constant of the perturbed motion, may be expressed to any order of ε . If J is used to eliminate z_1 , one winds up as required with a mechanical system containing only (n-2) independent variables.

An alternate method of deriving an invariant I for the system described by eq. (1) is based on an expansion of the Poisson bracket relation

$$\left[I, H \right] = 0 \tag{2}$$

This expansion has been described by Whittaker[2] and was further explored by McNamara and Whiteman [3] (henceforth referred to as McNW). The latter authors were able to show -- by a lengthy calculation and only to order ε^2 -- that the invariant thus obtained is, in a special case, related to Kruskal's J.

In what follows we shall show that for a system given by eq. (1), Kruskal's method may be modified to yield a family of invariants in a completely different way than that used in deriving J. It will then be shown that these invariants are solutions of eq. (2), are related to the invariant of McNW and also are connected with the adiabatic invariant J.

NOTATION

In order to obtain concise expressions, the notation used here departs somewhat from that of Kruskal. In Kruskal's work, $\underline{y}$ and $\underline{z}$ have $(n-1)$ components and special symbols ν and ϕ stand for what we here denote by y_n and z_n. In what follows, such vectors with $(n-1)$ components, excluding the angle variable, will be denoted by a tilde, e.g. $\underline{\tilde{y}}$, $\underline{\tilde{z}}$. Furthermore, if the canonical variables in $\underline{y}$ are arranged in order

$$\underline{y} = (\underline{p} , \underline{q}) \tag{3}$$

we shall define a conjugate vector

$$\underline{\bar{y}} = (\underline{q} , -\underline{p}) \tag{4}$$

so that $\bar{y}_1 = y_n$, $\bar{y}_n = -y_1$.

The use of $\bar{\underline{y}}$ enables one to write Poisson brackets concisely (the summation convention is henceforth used in all summations over n or (n-1) components) as

$$[a, b] = (\partial a/\partial \bar{y}_i)(\partial b/\partial y_i) \tag{5}$$

and Hamilton's equations become

$$dy_i/dt = -\partial H/\partial \bar{y}_i \tag{6}$$

Finally, we shall assume (as in Kruskal's work) that the basic period in the dependence on y_n is unity and denote quantities averaged over y_n (which clearly depend on $\tilde{\underline{y}}$ only) by angular brackets

$$\langle a \rangle = \int_0^1 a \; dy_n$$

KRUSKAL'S EXPANSION

Kruskal's method does not require the system to be canonical but assumes the evolution of $\underline{y}$ to be given by equations of the form

$$d\underline{y} / dt = \sum_{k=0}^{\infty} \varepsilon^k \underline{g}^{(k)}(\underline{y}) \tag{7}$$

where the components of $\underline{g}^{(k)}$ are periodic in y_n with period unity and where $\underline{g}^{(0)}$ has only one non-zero component, namely the n-th. Since we are interested in systems for which (7) reduces to (6) with H given by (1), we shall assume that this component is unity

$$\underline{g}^{(0)} = (0, 0 \cdots 0, 1) \tag{8}$$

although what follows can be extended to more general cases. We now seek a near-identity transformation to new variables

$$\underline{z} = \underline{y} + \sum_{k=1} \varepsilon^k \underline{\zeta}^{(k)}(\underline{y}) \tag{9}$$

such that the evolution of $\underline{z}$ satisfies

$$d\underline{z} / dt = \sum_{k=0} \varepsilon^k \underline{h}^{(k)}(\tilde{\underline{z}}) \tag{10}$$

with $h^{(0)} = g^{(0)}$. Since z_n does not appear on the right-hand side, the first $(n-1)$ equations of (10) form an autonomous system which may be solved independently, the solution then being substituted in the remaining equation to provide the evolution of z_n .

Substitution of (9) in the l.h.s. of (10) gives

$$\begin{aligned} d\underline{z} / dt &= d\underline{y} / dt + \sum_{k=1} \varepsilon^k (\partial \underline{\zeta}^{(k)} / \partial y_i)(dy_i / dt) \qquad (11) \\ &= \sum_{k=0} \varepsilon^k \left\{ \underline{g}^{(k)} + \sum_{j=0}^{k-1} g_i^{(j)} (\partial \underline{\zeta}^{(k-j)} / \partial y_i) \right\} \\ &= \sum_{k=0} \varepsilon^k \left\{ \underline{g}^{(k)} + (\partial \underline{\zeta}^{(k)} / \partial y_n) + \sum_{j=1}^{k-1} \underline{g}^{(j)} \cdot \nabla \underline{\zeta}^{(k-j)} \right\} \end{aligned}$$

with the ∇ operator defined in $\underline{y}$ space. We now convert the r.h.s. of (10) to depend on $\underline{y}$ as well, using the Taylor expansion operators derived as follows. Defining the exponential of a differential operator by means of the power series for e^x , one can formally express the Taylor expansion of $\underline{h}^{(k)}(\tilde{\underline{z}})$ as follows ($*$ stands for "operates on")

$$\begin{aligned} \underline{h}^{(k)}(\tilde{\underline{z}}) &= \underline{h}^{(k)}(\tilde{\underline{y}} + \sum \varepsilon^j \tilde{\underline{\zeta}}^{(j)}) \\ &= \exp (\sum \varepsilon^j \tilde{\underline{\zeta}}^{(j)} \cdot \nabla) * \underline{h}^{(k)}(\tilde{\underline{y}}) \\ &= \sum_{j=0} \varepsilon^j T^{(j)} * \underline{h}^{(k)}(\tilde{\underline{y}}) \end{aligned} \tag{12}$$

where the $T^{(j)}$ are differential operators obtained by expanding the exponential. The first few of them are [4]

$$
\begin{aligned}
T^{(0)} &= 1 \\
T^{(1)} &= \underline{\zeta}^{(1)}\cdot\nabla \\
T^{(2)} &= \underline{\zeta}^{(2)}\cdot\nabla + \tfrac{1}{2}\,\underline{\zeta}^{(1)}\underline{\zeta}^{(1)}:\nabla\nabla \qquad (13)\\
T^{(3)} &= \underline{\zeta}^{(3)}\cdot\nabla + \underline{\zeta}^{(1)}\underline{\zeta}^{(2)}:\nabla\nabla + (1/6)\,\underline{\zeta}^{(1)}\underline{\zeta}^{(1)}\underline{\zeta}^{(1)}\therefore\nabla\nabla\nabla
\end{aligned}
$$

Because $T^{(0)}$ equals unity, one can separate the $\underline{h}^{(k)}$ term from the rest. The r.h.s. of (10) then becomes

$$\sum \varepsilon^{k}\left\{\underline{h}^{(k)}(\tilde{\underline{y}}) + \sum_{j=1}^{k-1} T^{(j)} * \underline{h}^{(k-j)}(\tilde{\underline{y}})\right\} \qquad (14)$$

Both sides of (10) are now functions of $\underline{y}$ and therefore the equality holds independently for every order. One thus obtains a set of equations

$$\partial\underline{\zeta}^{(k)}/\partial y_n - \underline{h}^{(k)}(\tilde{\underline{y}}) = \underline{\lambda}^{(k)}(\underline{y}) \qquad (15)$$

where

$$\underline{\lambda}^{(k)}(\underline{y}) = \sum_{j=1}^{k-1} T^{(j)} * \underline{h}^{(k-j)} - \underline{g}^{(k)} - \sum_{j=1}^{k-1} \underline{g}^{(j)}\cdot\nabla\underline{\zeta}^{(k-j)} \qquad (16)$$

and

$$k = 1, 2, \ldots$$

This may be used as a recursion relation. Suppose all quantities entering here are (like the $\underline{g}^{(k)}$) either periodic in y_n or independent of it. Then $\underline{\lambda}^{(k)}$ will possess a "secular" part $\langle\underline{\lambda}^{(k)}\rangle$ independent of y_n and a purely periodic part averaging zero

$$(\underline{\lambda}^{(k)})_{per} = \underline{\lambda}^{(k)} - \langle\underline{\lambda}^{(k)}\rangle$$

On the l.h.s. of (15), $\underline{h}^{(k)}$ is wholly secular by definition, while the other term there is purely periodic, since any secular part of $\underline{\zeta}^{(k)}$ is removed by the differentiation. Thus if $\underline{\lambda}^{(k)}$ is given we can derive the k-th order quantities through

$$\underline{h}^{(k)}(\tilde{\underline{y}}) = -\langle \lambda^{(k)} \rangle \tag{17}$$

$$\partial \underline{\zeta}^{(k)}/\partial y_n = \underline{\lambda}^{(k)} - \langle \underline{\lambda}^{(k)} \rangle \tag{18}$$

from which

$$\underline{\zeta}^{(k)} = \int_0^{y_n} \left\{ \underline{\lambda}^{(k)} - \langle \underline{\lambda}^{(k)} \rangle \right\} dy_n' + \underline{\mu}^{(k)}(\tilde{\underline{y}}) \tag{19}$$

with $\underline{\mu}^{(k)}$ an arbitrary secular vector. In Kruskal's work, $\underline{\mu}^{(k)}$ is chosen to vanish, so that at $y_n = 0$ the vector $\underline{z}$ is identical with $\underline{y}$; we shall denote the functions thus obtained by $\hat{\underline{\zeta}}^{(k)}$. If however we only demand that $\underline{z}$ should satisfy equations of the form (10), $\underline{\mu}^{(k)}$ may be arbitrarily chosen. In what follows we shall make use of this free choice in order to endow the "nice variables" $\underline{z}$ with additional desired properties.

THE NEW INVARIANT

So far we have treated the general case of Kruskal's expansion (apart from our choice of $g_n^{(0)}$) with no reference to the canonical character of $\underline{y}$. Taking this now into account, one finds from (1) and (6)

$$g_i^{(k)} = -\partial H^{(k)}/\partial \bar{y}_i \tag{20}$$

Defining for convenience

$$\underline{\zeta}^{(0)} \equiv \underline{y}$$

and using (16) and (5), gives

$$\underset{\sim}{\lambda}^{(k)} = \sum_{j=1}^{k-1} \left\{ T^{(j)} * \underset{\sim}{h}^{(k-j)}(\tilde{\underset{\sim}{y}}) \right\} + \sum_{j=1}^{k} \left[H^{(j)}, \underset{\sim}{\zeta}^{(k-j)} \right] \qquad (21)$$

We now pose the following question: is it possible, by proper choice of the $\underset{\sim}{\mu}^{(k)}$ in (19), to make some component z_i of $\underset{\sim}{z}$ a constant Z_i of the motion?

If z_i is conserved, this means that $h_i^{(k)}$ vanishes for all k, which in turn implies the vanishing of $\langle \lambda_i^{(k)} \rangle$. By the last equation, this reduces to

$$\sum_{j=1}^{k} \left\langle \left[H^{(j)}, \zeta_i^{(k-j)} \right] \right\rangle = 0 \qquad (22)$$

Suppose that at the stage when eq. (22) is reached the $\underset{\sim}{\mu}^{(i)}$ have been derived up to and including the $(k-2)$ order. We can then fulfil (22) by choosing $\mu_i^{(k-1)}$ to satisfy

$$\left[\mu_i^{(k-1)}, \langle H^{(1)} \rangle \right] = \left\langle \left[H^{(1)}, \hat{\zeta}_i^{(k-1)} \right] \right\rangle + \sum_{j=2}^{k} \left\langle \left[H^{(j)}, \zeta_i^{(k-j)} \right] \right\rangle \qquad (23)$$

where the r.h.s. is assumed to be known at that stage. The above equation is a linear first-order partial differential equation, and solutions in general do exist. Deriving them explicitely is another problem, however. For the special case when all $H^{(k)}$ with $k > 1$ vanish, McNamara and Whiteman (who arrive at a similar equation) obtained from first principles formulas which allow $\mu_i^{(k)}$ to be derived (for $i=1$, which is the relevant case, as will be seen) up to $k=3$.

A different approach to the problem will be outlined in the last section of this article.

The iteration for Z_1 can thus be carried on -- provided it can be started. The first additive function encountered is $\mu_1^{(1)}$, used in ensuring the vanishing of $\langle \lambda_1^{(2)} \rangle$. There exists no adjustable variable to ensure the

the vanishing of $\langle \lambda_i^{(1)} \rangle$, so the iteration can be started if and only if this term vanishes of its own accord, which in turn implies

$$\langle \partial H^{(1)} / \partial \tilde{y}_i \rangle = 0 \tag{24}$$

If (as assumed) y_n enters H only through angular terms, this will certainly hold for i=1, since the y_n derivative which is applied in that case removes the secular part of $H^{(1)}$, leaving a purely periodic function. Thus an invariant Z_1 of the type discussed here may in general be derived. If $H^{(1)}$ itself is purely periodic, other invariants may be generated for $i \neq 1$.

THE POISSON BRACKET METHOD

McNamara and Whiteman[3], following Whittaker[2], derive an invariant I (in their notation: J) in the following way. Let I have an expansion in ε

$$I = \sum_{k=0} \varepsilon^k I^{(k)}(\underline{y}) \tag{25}$$

and let H be expanded as in (1) (this is a slight generalization: in the cited work, all $H^{(i)}$ with $i > 1$ vanish[5]). Then

$$[I, H] = 0 = \sum_{k=0} \varepsilon^k \sum_{j=0}^{k} [I^{(j)}, H^{(k-j)}] \tag{26}$$

By (5)

$$[I^{(k)}, H^{(0)}] = [I^{(k)}, y_1] = \partial I^{(k)} / \partial y_n \tag{27}$$

So the equation for the k-th order is

$$\partial I^{(k)} / \partial y_n = - \sum_{j=1}^{k-1} [I^{(j)}, H^{(k-j)}] \tag{28}$$

which defines $I^{(k)}$ recursively within an arbitrary function of $\tilde{\underline{y}}$

$$I^{(k)} = -\int_0^{y_n} \sum_{j=0}^{k-1} \left[I^{(j)}, H^{(k-j)} \right] dy_n' + G^{(k)}(\tilde{y}) \tag{29}$$

$$= \hat{I}^{(k)} + G^{(k)}$$

By (19) and (21), this is exactly the same as the equation for $\zeta_i^{(k)}$, provided all $h_i^{(k)}$ vanish.

Consider again eq. (28) : due to the y_n derivative, its l.h.s. will be purely periodic, but unless special steps are taken the r.h.s. may well contain a secular part. We therefore must assume that at the stage at which $I^{(k)}$ is being derived, $G^{(k-1)}$ has not yet been determined. The r.h.s. can then be made purely periodic by requiring

$$\left[G^{(k-1)}, \langle H^{(1)} \rangle \right] = -\left\langle \left[\hat{I}^{(k-1)}, H^{(1)} \right] \right\rangle - \sum_{j=0}^{k-2} \left\langle \left[I^{(j)}, H^{(k-j)} \right] \right\rangle \tag{30}$$

which has the same form as (23).

The lowest order of (26) gives

$$\left[I^{(0)}, H^{(0)} \right] = 0 = \partial I^{(0)} / \partial y_n$$

so that

$$I^{(0)} = I^{(0)}(\tilde{y})$$

The iteration can be started only if $I^{(0)}$ makes the r.h.s of (28) purely periodic for $k=1$:

$$\left[I^{(0)}, \langle H^{(1)} \rangle \right] = 0 \tag{31}$$

Obviously, any function of $H^{(1)}$ can be chosen as $I^{(0)}$; McNW choose

$$I^{(0)} = \langle H^{(1)} \rangle \tag{32}$$

This however is not entirely satisfactory, since we expect $I^{(0)}$ to tend to some natural invariant of the unperturbed system in the limit of vanishing ε , independent of the choice of $H^{(1)}$. A more suitable choice is

$$I^{(0)} = y_1 \tag{33}$$

which also satisfies (31), since (24) holds for $i=1$. With this choice, I equals the action variable y_1 in the unperturbed limit, a property shared by the invariant Z_1 previously derived and also (it may be shown) by Kruskal's J . The alternative choice, made by McNW, will be explored in the next section.

With $I^{(0)}$ chosen as in (33), it also equals $\zeta_1^{(0)}$ and it is easy to show that the expansion equations of I match those of the invariant Z_1 stage by stage. On may then match

$$I^{(k)} = \zeta_1^{(k)}$$

$$G^{(k)} = \mu_1^{(k)}$$

by making identical choices of the arbitrary functions of $H^{(1)}$ (and by virtue of (33) satisfying (31), of y_1 as well) which can be added to $G^{(k)}$ and to $\mu_1^{(k)}$ at every stage.

<u>THE CHOICE $I^{(0)} = \langle H^{(1)} \rangle$</u>

McNamara and Whiteman chose $I^{(0)}$ as in (32), in the special case where $H^{(k)}$ vanishes for $k>1$. In that case, their recursion continues with

$$I^{(1)} = - \int_0^{y_n} \left[\langle H^{(1)} \rangle , H^{(1)} \right] dy_n' + G^{(1)} \tag{34}$$

$$= \hat{I}^{(1)} + G^{(1)}$$

and

$$\left[G^{(1)},\ \langle H^{(1)}\rangle\right] = \left\langle \left[H^{(1)},\ \hat{I}^{(1)}\right]\right\rangle \qquad (35)$$

Expanding the invariant Z_1 by Kruskal's method, for the same H,

$$\lambda_1^{(1)} = \partial H^{(1)}/\partial y_n$$

and by (19)

$$\zeta_1^{(1)} = H^{(1)} - H^{(1)}(y_n=0) + \mu_1^{(1)}$$

From (23) and the preceding equation, noting that secular functions may be taken out of the averaging brackets

$$\left[\mu_1^{(1)},\ \langle H^{(1)}\rangle\right] = -\left[\langle H^{(1)}\rangle,\ H^{(1)}(y_n=0)\right]$$

Let us select

$$\mu_1^{(1)} = -\langle H^{(1)}\rangle + H^{(1)}(y_n=0)$$

so that

$$\zeta_1^{(1)} = H^{(1)} - \langle H^{(1)}\rangle$$

The second order equations then are , by (21)

$$\zeta_1^{(2)} = -\int_0^{y_n}\left[H^{(1)},\ \langle H^{(1)}\rangle\right]dy_n' + \mu_1^{(2)}$$

and

$$\left[\mu_1^{(2)},\ \langle H^{(1)}\rangle\right] = \left\langle\left[H^{(1)},\ \hat{\zeta}_1^{(2)}\right]\right\rangle$$

This demonstrates that, with matching choices of additive functions

$$I^{(1)} = -\zeta_1^{(2)}$$

$$G^{(1)} = -\mu_1^{(2)}$$

If additive functions of higher orders are matched as well, the subsequent parts of $-\underline{\zeta}_1$ and of I are identical, except for an extra order of ε in the former. That means

$$- \varepsilon (I - I^{(0)}) = Z_1 - \zeta_1^{(0)} - \varepsilon \zeta_1^{(1)}$$

or

$$Z_1 = H - \varepsilon I \qquad (36)$$

a relation resembling that obtained (to order ε^2) by McNW , except that in their result J appears in place of Z_1 .

<u>RELATION TO J</u>

Kruskal defined the adiabatic invariant J by

$$J = \oint \sum_k p_k \, dq_k = \int_0^1 \sum_k p_k \, (\partial q_k / \partial z_n) \, dz_n \qquad (37)$$

where the integration is carried over a set of points ("ring") sharing the same $\underline{z}$ and differing only in z_n . For details about J , the reader is referred to Kruskal's article[1]; its value is independent of the canonical set used in its derivation, though the components of $\underline{y}$ are the best choice for this role, if they form a canonical set (the inverse transformation $\underline{z} \rightarrow \underline{y}$ must then be derived by means of the expansion operators (13)). Here we shall merely sketch out the connection between Z_1 and J without deriving the details.

Suppose that among the many sets of "nice" variables possible, differing in their choices of $\mu_i^{(k)}$ but all obeying (10), there exists a set (or a family of sets) that is canonical, with z_1 conjugate to z_n . Obviously, this set, too, can be used in deriving (37), leading immediately to

$$J = z_1$$

The existence of nice canonical variables has been proved by Kruskal[1] in appendix 2 of his article. It is furthermore possible to express the $\mu_i^{(k)}$ which generate such variables. The details of this derivation are somewhat involved and will therefore be described in a separate article; here we shall just assume that these $\mu_i^{(k)}$ are known. Then the transformation which they describe belongs to the (much larger) family of transformations which make z_1 a constant of the motion, each of which in its turn provides a solution to the corresponding Poisson bracket expansion. Thus among the many possible solutions of Z_1 and I, there exist such ones for which

$$I = Z_1 = J$$

Practically, given the $\mu_i^{(k)}$ which make $\underline{z}$ canonical, these functions probably offer the best way of deriving Z_1 or I, since they are known to solve eq. (30). The only other possibility for solving this equation is to use the formulas of McNW, which are valid for the lowest few orders only and are specifically related to the choice of $I^{(0)}$ given in (32).

References and Comments

(1) M. Kruskal, J. Math. Physics 3, 806 (1962)

(2) E.T. Whittaker, Analytical Dynamics of Particles and Rigid Bodies, 4th Edition, Dover 1944.

(3) B. McNamara and K.J. Whiteman, J. Math. Physics 8, 2029 (1967)

(4) These operators originated in the calculation of derivatives of a function of a function by Faa De Bruno, Quarterly J. of Math. 1, 359 (1855). They are listed up to $j = 8$ by P. Musen, J. of Astronautical Sciences 12, 129 (1965).

(5) The definition of Poisson brackets by McNamara and Whiteman, in eq. 3-2 of their article, has a reversed sign compared to the usual convention. As a consequence, some equations of McNW differ in sign from those obtained here.