X-641-69-516

PREPRINT

NASA TM X-63848

DIRECT CANONICAL TRANSFORMATIONS

DAVID P. STERN

DECEMBER 1969
REVISED
MARCH 1970

GSFC — GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

Direct Canonical Transformations

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

Abstract

Some of the perturbation methods in classical Hamiltonian mechanics lead to near-identity transformations of the variables, with the new variables explicitely given as functions of the old ones. Two methods are used for identifying and characterizing the sub-class of all such transformations which are also canonical; one approach is related to the conventional method of generating canonical transformations, while the other one uses the properties of Poisson brackets and is related to an operator method of Lie. Either of the methods may be used to derive certain steps in a perturbation method devised by Lacina, inadvertedly omitted by that author.

INTRODUCTION

Let a classical canonical system be given, described by $2n$ canonical variables collectively denoted by the vector $\underline{y}$. The components y_i may be divided into canonical momenta p_i and canonical coordinates q_i, and we shall assume that their order is

$$\begin{aligned} \underline{y} &= (p_1 \ldots p_n , q_1 \ldots q_n) \\ &= (\underline{p} , \underline{q}) \end{aligned} \tag{1}$$

A transformation $\underline{y} \rightarrow \underline{z}$ is termed canonical if the new variables also form a canonical set

$$\underline{z} = (\underline{P} , \underline{Q})$$

Canonical transformations are customarily defined by means of a generating function σ (1)(2) depending on n new and n old variables (and possibly on time, though this will not be assumed in what follows). Of the several possible types of such functions, we shall choose to use those of the form

$$\sigma = \sigma(\underline{P} , \underline{q}) \tag{2}$$

for which the transformation equations are given by

$$\begin{aligned} \partial\sigma/\partial P_i &= Q_i \\ \partial\sigma/\partial q_i &= p_i \end{aligned} \tag{3}$$

These equations give the transformations indirectly: if the old variables are given and the new ones are sought, additional "untangling" is generally in order. By contrast, a "direct" transformation

$$\underline{z} = \underline{z}(\underline{y})$$

is immediately usable. In particular, we shall be concerned with near-identity transformations expanded in terms of a small parameter $\varepsilon \ll 1$ (bracketed superscript denotes order in ε):

$$\underline{z} = \underline{y} + \sum_{k=1} \varepsilon^{k} \underline{\zeta}^{(k)}(\underline{y}) \qquad (4)$$

In what follows, we shall investigate the conditions under which the transformation (4) is canonical.

Although not essential to what follows, the motivation for this investigation may be of some interest. In the classical perturbation theory developed for celestial mechanics, the basic problem is the solution of a canonical system depending on a small parameter ε, given that in the limit of vanishing ε ("unperturbed motion") the solution is known and periodic. One procedure used there (usually associated with the names of Von Zeipel and Poincare) is as follows[3]. First, one prepares the ground by transforming the original problem to new variables -- which will be labeled $\underline{y}$ as in eq.(1) -- so that one of the new variables is angle-like and represents the periodicity of the unperturbed motion, while the others, in the limit of vanishing ε, are constant in time. If the angle variable is conjugate to the variable y_1, the new variables may be chosen so as to make the hamiltonian H equal y_1 as well.

The same variables are then introduced into the finite-ε problem, in which case the Hamiltonian assumes the form

$$H = y_1 + \sum_{k=1} \varepsilon^{k} H^{(k)}(\underline{y}) \qquad (5)$$

A near-identity canonical transformation is now sought, generated by

$$\sigma(\underline{p}, \underline{q}) = \sum_{i=1} p_i q_i + \sum_{k=1} \varepsilon^k \sigma^{(k)}(\underline{p}, \underline{q}) \qquad (7)$$

such that in the new variables $\underline{z}$, the angle variable is absent from the transformed Hamiltonian , making its conjugate z_1 a constant of the motion. Further solution of the motion involves neither of these variables as an unknown, so that the problem has been reduced by two variables, or equivalently, by one dimension.

A different approach, independently developed by Kruskal[(4)], is related to the method of Bogoliubov and Krylov, and describes the evolution of a system through $2n$ equations of the form

$$d\underline{x}/dt = \sum_{k=0} \varepsilon^k \underline{f}^{(k)}(\underline{x}) \qquad (8)$$

The method does not require the system to be canonical, but we shall restrict ourselves here to problems where this is the case. The preparatory derivation of "intermediate variables" $\underline{y}$ is then very much the same as before. To solve the perturbed problem, however, the method seeks a direct transformation of the form (4) , with the angle variable transformed into one of the components of $\underline{z}$ which will be denoted (for reasons which will be clarified later) as $\bar{z}_1$. The transformation is such that when the equations (8) are transformed to give the evolution of $\underline{z}$

$$d\underline{z}/dt = \sum_{k=0} \varepsilon^k \underline{h}^{(k)}(\underline{z}) \qquad (9)$$

then $\bar{z}_1$ is absent on the right hand side. One can now separately handle the $(2n-1)$ equations in $(2n-1)$ variables obtained by omitting the equation for $d\bar{z}_1/dt$ from the above set. One can also obtain a constant of the motion, the so-called adiabatic invariant J , thus removing one more unknown variable from the problem.

The details of J will not be described here beyond noting that it resembles in many ways the constant action variable z_1 obtained by the preceding method -- for instance, the two may be shown to be equal at least to order ε^0

The two methods are evidently similar, and one may ask whether they can be made to coincide, by a suitable choice of the arbitrariness existing in Kruskal's method[5]. For this to happen, the transformation (4) obtained by Kruskal's approach must be canonical, which leads one to the basic problem stated before. As will be shown, the matter is also intimately connected to the work of Lacina[6],[7], where it leads to the correction of an error.

Two approaches to the problem will be described here: the first one is concise and is derived from the conventional form of canonical transformations, while the second one appears to be more elegant and has interesting geometrical implications. It is related to an operator method for characterising direct canonical transformations, originally devised by Lie[8] and recently applied to celestial mechanics by Hori[9] and Deprit[10].

DIRECT SOLUTION

If the transformation (4) is canonical, then there must exist a generating function $\sigma(\underline{P}, q)$ of the form given in (7) that leads to it. Applying (3), one obtains

$$Q_i = q_i + \sum_{k=1} \varepsilon^k \{ \partial\sigma^{(k)}(\underline{P}, \underline{q})/\partial P_i \} \qquad (10)$$

$$P_i = p_i - \sum_{k=1} \varepsilon^k \{ \partial\sigma^{(k)}(\underline{P}, \underline{q})/\partial q_i \} \qquad (11)$$

If the ordering of the components of $\underline{\zeta}^{(k)}$ corresponds to the ordering of the components of $\underline{y}$ in (1), one may define "partial vectors", the sum of which equals $\underline{\zeta}^{(k)}$

$$\underline{\pi}^{(k)} = (\zeta_1^{(k)} \ldots \zeta_n^{(k)}, 0 \ldots 0)$$
$$\underline{\theta}^{(k)} = (0 \ldots 0 , \zeta_{n+1}^{(k)} \ldots \zeta_{2n}^{(k)}) \tag{12}$$

which allows (4) to be split up into

$$Q_i = q_i + \sum_{k=1} \varepsilon^k \theta_{i+n}^{(k)}(\underline{y}) \tag{13}$$

$$P_i = p_i + \sum_{k=1} \varepsilon^k \pi_i^{(k)}(\underline{y}) \tag{14}$$

$$(i = 1 \ldots n)$$

Substituting in (11)

$$P_i = p_i - \sum_{k=1} \varepsilon^k \partial\sigma^{(k)}\left[\underline{p} + \sum_{m=1} \varepsilon^m \underline{\pi}^{(m)}(\underline{y}), \underline{q}\right] / \partial q_i \tag{15}$$

We now introduce expansion operators [(11)(5)] $S^{(k)}$ such that for any function f of the canonical variables ($*$ means 'operates on')

$$f(\underline{P}, \underline{q}) = \sum_{m=0} \varepsilon^m S^{(m)} * f(\underline{p}, \underline{q}) \tag{16}$$

The explicit expressions for $S^{(m)}$ may be obtained by replacing $\underline{\zeta}^{(k)}$ with $\underline{\pi}^{(k)}$ in equations (13) of reference 5 ; they involve ∇ operators in $\underline{y}$ space acting on the $\underline{\pi}^{(k)}$. We then get ($S^{(0)}$ being unity)

$$P_i = p_i - \sum_{k=1} \varepsilon^k \left\{ \partial\sigma^{(k)}/\partial q_i + \sum_{m=1}^{k-1} S^{(m)} * \partial\sigma^{(k-m)}/\partial q_i \right\} \tag{17}$$

where the various orders of σ on the r.h.s. are functions of $\underline{y}$ only and may be obtained from those appearing in eq. (7) by replacing $\underline{P}$ everywhere with $\underline{p}$. This is formally equivalent to (14) and therefore

$$\pi_i^{(k)} = - \partial\sigma^{(k)}(\underline{y})/\partial q_i - \sum_{m=1}^{k-1} S^{(m)} * \partial\sigma^{(k-m)}/\partial q_i \tag{18}$$

Since the $S^{(m)}$ appearing here all involve lower orders of $\underline{\pi}^{(m)}$, this ia a usable recursion formula, allowing $\underline{\pi}^{(k)}$ for the direct transformation to be solved -- provided the $\sigma^{(k)}$ are known and provided the lower orders have already been solved. Similarly, one gets

$$\theta_{i+n}^{(k)} = \partial\sigma^{(k)}(\underline{y})/\partial p_i + \sum_{m=1}^{k-1} S^{(m)}{}_* \ \partial\sigma^{(k-m)}/\partial p_i \tag{19}$$

The last two equations can be joined together by introducing the concept of the conjugate vector $\bar{\underline{y}}$, formed by permuting the order of components of $\underline{y}$ given in equation (1) to

$$\bar{\underline{y}} = (\underline{q}, -\underline{p}) \tag{20}$$

Several properties of $\bar{\underline{y}}$ are described in reference 5 ; using them, one may express Poisson brackets as ($\bar{\nabla}$ operator defined in $\bar{\underline{y}}$ space)

$$[a, b] = \bar{\nabla} a \cdot \nabla b \tag{21}$$

and in particular

$$[a, y_i] = \partial a/\partial \bar{y}_i \tag{22}$$

With this notation, (18) and (19) may be combined to

$$\underline{\zeta}^{(k)} = -\bar{\nabla}\sigma^{(k)} + \underline{\rho}^{(k)} \tag{23}$$

with

$$\rho_i^{(k)} = -\sum_{m=1}^{k-1} S^{(m)}{}_* \ \partial\sigma^{(k-m)}/\partial\bar{y}_i \tag{24}$$

being determined only by lower orders. In spite of its external appearance, this relation is not free from partial vectors, since the $\underline{\pi}^{(m)}$ appear in the S operators. A criterion for canonicity is now easily established. If the transformation (4) is canonical up to and including the order (k-1), one may form the vector

$$\underline{u}^{(k)} = \underline{\zeta}^{(k)} - \underline{\rho}^{(k)}$$

and the transformation will be canonical if and only if $\underline{u}^{(k)}$ is a gradient in $\underline{y}$ space, that is, if the components of the "conjugate curl" tensor of $\underline{u}^{(k)}$

$$(\bar{\nabla} \times \underline{u}^{(k)})_{ij} = \partial u_i^{(k)}/\partial \bar{y}_j - \partial u_j^{(k)}/\partial \bar{y}_i \tag{25}$$

all vanish.

SOLUTION BY POISSON BRACKETS

A different approach involves no partial vectors and is based on the properties of Poisson brackets. If both $\underline{y}$ and $\underline{z}$ are canonical, and equal in the zeroth order, then

$$[y_i, y_j] = [z_i, z_j] \tag{26}$$

the value of the brackets being 1 or 0 depending on whether canonical conjugacy exists or not. Substituting the expansion (4) on the r.h.s. and equating each order in ϵ separately, one gets as the condition for canonicity a set of relations, which through (22) and (25) can be expressed as

$$(\bar{\nabla} \times \underline{\zeta}^{(k)})_{ij} = - \sum_{m=1}^{k-1} [\zeta_i^{(m)}, \zeta_j^{(k-m)}] \tag{27}$$

Of more interest than a criterion for canonicity would be the derivation of a method for actually generating direct canonical transformations. Equations (27) are not very convenient for this, since they involve the components of $\underline{\zeta}^{(k)}$ only through combinations of their derivatives. A method for integrating these equations would obviously be useful here. For instance, the first equation of the set

$$\bar{\nabla} \times \underline{\zeta}^{(1)} = 0$$

may be integrated to

$$\underline{\zeta}^{(1)} = \bar{\nabla} \chi^{(1)} \tag{28}$$

with $\chi^{(1)}$ an arbitrary scalar. Similarly, an arbitrary conjugate gradient may be added to any of the $\underline{\zeta}^{(k)}$, since such a gradient is ignored by the curl operation. We may thus formally write the most general solution of (27) as

$$\underline{\zeta}^{(k)} = \bar{\nabla}\chi^{(k)} + \underline{f}^{(k)} \tag{29}$$

where $\underline{f}^{(k)}$ is any vector depending only on lower orders of $\underline{\zeta}^{(m)}$ which gives one particular solution of that equation. Comparing (29) with (25) shows that one possible choice of $\underline{f}^{(k)}$ is

$$\underline{f}^{(k)} = \underline{\xi}^{(k)} \tag{30}$$

which carries with it the identification

$$\chi^{(k)} = -\sigma^{(k)}$$

The drawback of this choice (more esthetical than practical) is that $\underline{f}^{(k)}$ contains partial vectors. Solutions which do not split up phase space into coordinates and momenta also do exist, to any order. The first of these is almost trivial: by (28), $\underline{f}^{(1)}$ vanishes. The next three vectors are

$$\begin{aligned}
\underline{f}^{(2)} &= \tfrac{1}{2}\ \underline{\zeta}^{(1)}\cdot\nabla\underline{\zeta}^{(1)} \\
\underline{f}^{(3)} &= \underline{\zeta}^{(2)}\cdot\nabla\underline{\zeta}^{(1)} \\
\underline{f}^{(4)} &= \underline{\zeta}^{(3)}\cdot\nabla\underline{\zeta}^{(1)} + \tfrac{1}{2}\underline{\zeta}^{(2)}\cdot\nabla\underline{\zeta}^{(2)} \\
&\quad + \tfrac{1}{4}\left\{\left(\underline{\zeta}^{(1)}\cdot\nabla\underline{\zeta}^{(1)}\right)\cdot\nabla\underline{\zeta}^{(2)} - \underline{\zeta}^{(2)}\cdot\nabla\left(\underline{\zeta}^{(1)}\cdot\nabla\underline{\zeta}^{(1)}\right)\right\} \\
&\quad - \tfrac{1}{4}\left[\left(\underline{\zeta}^{(1)}\cdot\nabla\underline{\zeta}^{(1)}\right)\cdot\nabla\underline{\zeta}^{(1)}\right]\cdot\nabla\underline{\zeta}^{(1)}
\end{aligned} \tag{31}$$

They are far from unique, since various "curl-free" expressions, involving only the $\underline{\zeta}^{(m)}$, may be added to any one of them. A general method for deriving such expressions, based on Lie's method for characterizing direct canonical transformations, will now be described.

LIE'S METHOD

Let $W(\underline{y})$ be an arbitrary function of a given set of canonical variables. One then defines[10] the Lie derivative generated by W of any function $f(\underline{y})$ as the function

$$L_W(f) = [\, f, W\,]$$

The operator L_W is linear

$$L_W(\alpha f + \beta g) = \alpha L_W(f) + \beta L_W(g)$$

and its action on a product resembles that of the derivative

$$L_W(fg) = f L_W(g) + g L_W(f)$$

By means of Jacobi's identity one can prove similar properties for L_W acting on Poisson brackets, except that the order of terms must now be preserved:

$$L_W [f, g] = [L_W(f), g] + [f, L_W(g)]$$

By successive application of L_W, various powers $L_W^{\,n}$ of the operator may be defined, and for completeness one then includes $L_W^{\,0}$ as the identity operator. If α is a constant, $L_W^{\,n}(\alpha)$ thus vanishes for all values of n except zero, since all powers of L_W except the zeroth involve differentiation.

Using the preceding definitions, one may define an exponential operator

$$\exp(\varepsilon L_W) * f = (1 + \varepsilon L_W + \tfrac{1}{2}\varepsilon^2 L_W^{\,2} + \ldots) * f \qquad (32)$$

where ε is a constant much smaller than unity which helps the expression converge (convergence will not be discussed here, however). The property of the exponential operator of importance here is [10]

$$\exp(\varepsilon L_W) * [f, g] = [\exp(\varepsilon L_W) * f \;,\; \exp(\varepsilon L_W) * g]$$

The preceding transformation shows that if $\underline{y}$ is canonical, then for any W

$$\underline{z} = \exp \varepsilon L_W * \underline{y} \tag{33}$$

is also canonical, for we then have

$$\begin{aligned}[z_i, z_j] &= \exp \varepsilon L_W * [y_i, y_j] \\ &= [y_i, y_j]\end{aligned}$$

the latter Poisson brackets being always a constant, equal to 1 or 0. By (32), the transformation is a near-identity one, even if the "generating function" W does not depend on ε. It is, nevertheless, possible to include such a dependence , expressed in a power series in ε

$$W(\underline{y}, \varepsilon) = \sum_{k=0} \varepsilon^k W^{(k)}(\underline{y}) \tag{34}$$

We now prove the following. Let a direct canonical transformation be given by Lie's method as in (33), with W expanded as in (34). Then the same transformation may also be expressed as in (4), with $\underline{\zeta}^{(k)}$ given as in (29). If we choose to identify

$$\chi^{(k)} = - W^{(k-1)} \tag{35}$$

then the two approaches may be readily related and an explicit expression for $\underline{f}^{(k)}$ may be found.

We begin with the following lemma: suppose that (35) holds, that $i < k$, that the "function" $g(\underline{\zeta})$ is an expression involving orders of $\underline{\zeta}^{(m)}$ lower than the k-th and that $f^{(m)}(\underline{\zeta})$ is known for values of m smaller than k. Then a "function" $h(\underline{\zeta})$, similar in structure to g and $f^{(m)}$, may always be found so that

$$\left[g(\underline{\zeta}), W^{i-1} \right] = h(\underline{\zeta})$$

Proof: using (21), (29) and (35)

$$\begin{aligned}\left[g, W^{i-1} \right] &= \left[\chi^{(i)}, g \right] \\ &= \bar{\nabla} \chi^{(i)} \cdot \nabla g(\underline{\zeta}) \\ &= \left(\underline{\zeta}^{(i)} - \underline{f}^{(i)}(\underline{\zeta}) \right) \cdot \nabla g(\underline{\zeta})\end{aligned}$$

and the last expression has the required form.

Several corrolaries are now easily derived:

(1) The preceding result is still valid if g is replaced by a vector $\underline{g}(\underline{\zeta})$ in $\underline{y}$ space, in which case h is also replaced by a vector $\underline{h}$.

(2) If multiple nested Poisson brackets are given, under the same conditions as stated before and of the form

$$\left[\dots \left[\left[\underline{g}(\underline{\zeta}), W^{i-1} \right], W^{j-1} \right], \dots W^{s-1} \right] \tag{36}$$

they may still be reduced to the form $\underline{h}(\underline{\zeta})$. To prove this one only has to note that the innermost brackets may be thus reduced, then the innermost brackets of the remaining expression, and so on until all brackets have been eliminated.

(3) The preceding still holds if $\underline{g}(\underline{\zeta})$ in (36) is replaced by $\underline{y}$. For the innermost brackets then become

$$\begin{aligned}\left[\underline{y}, W^{i-1} \right] &= \left[\chi^{(i)}, \underline{y} \right] \\ &= \bar{\nabla} \chi^{(i)} \\ &= \underline{\zeta}^{(i)} - \underline{f}^{(i)}(\underline{\zeta})\end{aligned}$$

which has the form of $\underline{h}(\underline{\zeta})$. The remaining brackets may then be removed as before.

We now prove the main assertion. Let

$$L_W^{(k)}(\underline{g}) \quad = \quad \left[\, \underline{g},\ W^{(k-1)} \right]$$

(note displacement of order index) so that

$$\varepsilon\, L_W \quad = \quad \sum_{k=1} \varepsilon^k\, L_W^{(k)}$$

With this notation, the expression (36) may be rewritten

$$L_W^{(s)} \ldots L_W^{(j)}\, L_W^{(i)}(\underline{g}) \tag{37}$$

and as was noted, this can be reduced to the form $\underline{h}(\underline{\xi})$, as can the analogous expression with $\underline{y}$ replacing $\underline{g}$.

We now have, by (32)

$$\underline{\underline{z}} \quad = \quad \left\{ 1 + \sum \varepsilon^k\, L_W^{(k)} + \tfrac{1}{2} \left(\sum \varepsilon^k\, L_W^{(k)} \right)^2 + \ldots \right\} * \underline{y}$$

$$\sum_{k=0} \varepsilon^k\, M^{(k)} * \underline{y} \tag{38}$$

where the $M^{(k)}$ are expansion operators resembling those of (16), with the difference that account must be taken of the fact that $L_W^{(k)}$ operators with different values of k do not commute. One has, for instance (compare the last of eqs. (13) in ref. 5)

$$M^{(3)} \;=\; L_W^{(3)} + \tfrac{1}{2}\left(L_W^{(1)}\, L_W^{(2)} + L_W^{(2)}\, L_W^{(1)} \right) + (1/6)\, (L_W^{(1)})^3$$

Let us denote

$$N^{(k)} \;=\; M^{(k)} - L_W^{(k)}$$

Then, since

$$L_W^{(k)}(\underline{y}) \;=\; \left[\, \underline{y},\ W^{(k-1)} \right] = \; - \,\bar{\nabla}\, W^{(k-1)}$$

one obtains from (38)

$$\underline{\underline{z}} = \underline{y} - \sum_{k=1} \varepsilon^k\, \bar{\nabla}\, W^{(k-1)} + \sum_{k=1} \varepsilon^k\, N^{(k)} * \underline{y} \tag{39}$$

Since $N^{(k)} * \underset{\sim}{y}$ consists only of terms of the form (37), it may be reduced to a "function" $\underline{h}(\underset{\sim}{\zeta})$. Comparison of the last equation with (4) and (29) then identifies this function with $\underline{f}^{(k)}(\underset{\sim}{\zeta})$, provided (35) holds. This completes the proof of our original assertion.

LACINA'S EXPANSION

Lacina[6][7] has published what he claims is a simple new canonical perturbation method, leading to results similar to those obtained from the Hamilton-Jacobi equation. Unfortunately, the simplicity is more apparent than real, for two reasons. First, there exists an important omission in the calculation, and secondly, there is no assumption of near-periodicity, so that the elimination of secular terms may be dispensed with. Perturbation calculations for systems without periodic character are possible, but of little interest, since their range of validity in time is usually quite limited. It is the periodic character inherent in the problems of celestial mechanics and of guiding center motion which makes possible solutions valid over long intervals in time, provided secular terms are eliminated.

Lacina's result is easily derived by the preceding formalism and, in fact, our notation allows more concise treatment than is found in the original articles, which use a separate notation for canonical coordinates and momenta. Let a near-identity canonical transformation be given by a direct relation as in (4) ; the transformation is then fully specified by the various orders of $\chi^{(k)}$ appearing in (29), assuming of course that a paticular choice of $\underline{f}^{(k)}$ has been selected. This choice could be the one of (30), in which case $\chi^{(k)}$ is the k-th order of the conventional generating function σ , with sign reversed, or it may be the one derived in (39): any such choice may be used in what follows.

By using (29), any single component of the direct expansion, e.g.

$$\zeta_1^{(k)} = \partial \chi^{(k)}/\partial \bar{y}_1 + f_1^{(k)} \tag{40}$$

may be used to define $\chi^{(k)}$ (within a certain arbitrariness) and consequently the transformation. This is essentially the idea behind Lacina's approach. To define the transformation by means of $\zeta_1^{(k)}$, he uses the $(i, 1)$ component of (27)

$$\partial \zeta_i^{(k)}/\partial \bar{y}_1 = \partial \zeta_1^{(k)}/\partial \bar{y}_i - \sum_{m=1}^{k-1} \left[\zeta_i^{(m)}, \zeta_1^{(k-m)}\right] \tag{41}$$

from which

$$\zeta_i^{(k)} = \int_C^{\bar{y}_1} \left\{ \partial \zeta_1^{(k)}/\partial \bar{y}_i - \sum_{m=1}^{k-1} \left[\zeta_i^{(m)}, \zeta_1^{(k-m)}\right]\right\} d\bar{y}_1' + \mu_i^{(k)}(\underset{\sim}{\tilde{y}}) \tag{42}$$

where $\underset{\sim}{\tilde{y}}$ is the vector formed of the $(2n-1)$ components of $\underset{\sim}{y}$ excluding $\pm \bar{y}_1$ (the sign being adjusted so that this is a component of $\underset{\sim}{y}$).

The lower limit is arbitrary, but its choice affects $\mu_i^{(k)}$, which has to be chosen in a way assuring that (27) also holds when neither i nor j equals 1. To handle such cases, we express the Poisson brackets by means of $\underset{\sim}{f}^{(k)}$, which is known to be a solution for $\underset{\sim}{\zeta}^{(k)}$ in (27)

$$\zeta_i^{(k)} = \int_C^{\bar{y}_1} \left\{ \partial \zeta_1^{(k)}/\partial \bar{y}_i + \partial f_i^{(k)}/\partial \bar{y}_1 - \partial f_1^{(k)}/\partial \bar{y}_i \right\} d\bar{y}_1' + \mu_i^{(k)}(\underset{\sim}{\tilde{y}}) \tag{43}$$

As one forms the (i,j) element of the curl of the vector of which the above expression is one component, the contributions of $\zeta_1^{(k)}$ and of $f_1^{(k)}$ in the integrand cancel out and one is left with

$$(\bar{\nabla} \times \underline{\zeta}^{(k)})_{ij} = \int_C^{\bar{y}_1} (\partial/\partial \bar{y}_1')(\bar{\nabla} \times \underline{f}^{(k)})_{ij}\, d\bar{y}_1' + (\bar{\nabla} \times \underline{\mu}^{(k)})_{ij}$$

$$= (\bar{\nabla} \times \underline{f}^{(k)})_{ij} - (\bar{\nabla} \times \underline{f}^{(k)}(\bar{y}_1{=}C))_{ij} + (\bar{\nabla} \times \underline{\mu}^{(k)})_{ij} \tag{44}$$

Here the l.h.s. cancels with the first term on the right and one is left with the condition

$$\underline{\mu}^{(k)}(\underline{\tilde{y}}) = \underline{f}^{(k)}(\bar{y}_1{=}C) + \bar{\nabla}\psi^{(k)}(\underline{\tilde{y}}) \tag{45}$$

where $\psi^{(k)}$ is any function of $\underline{\tilde{y}}$ (or zero). Because only the curl of $\underline{f}^{(k)}$ is involved, any choice satisfying (29) may be used. None of this appears in Lacina's work[7], because the contributions of the lower limit of integration are inadvertently omitted in the last steps of the equations following eq. (15) and (16) there.

Let us assume that the unperturbed problem has been prepared so that the Hamiltonian has the form given in eq. (5). Following Lacina, we now stipulate that in the new variables the Hamiltonian equals z_1. Since the transformation is time independent, this new Hamiltonian equals the old one, given in (5), leading to the identification

$$\zeta_1^{(k)}(\underline{y}) = H^{(k)}(\underline{y}) \tag{46}$$

Substituting this choice of $\zeta_1^{(k)}$ into (42) and evaluating $\zeta_i^{(k)}$ by means of (45) allows the other z_i to be derived to any desired order. Since the new Hamiltonian equals z_1, the variable conjugate to z_1 will be linear in time and all other new variables (z_1 included) constants of the motion. The problem is thus essentially solved.

If one defines

$$\zeta_i^{(0)} = y_i$$

and uses (42) and (22) in (41), then the basic equation converts to

$$\sum_{m=0}^{k} \left[\dot{\zeta}_i^{(m)} , H^{(k-m)} \right] = 0$$

This is the starting point of the method of McNamara and Whiteman [12][5], which in turn is related to Whittaker's adelphic integral[13], Contopoulos' third integral[14] and to other approaches quoted by Contopoulos. The difference is that McNamara and Whiteman assume a periodic character of the motion and include an extra step to ensure elimination of secular terms. While their aim is to generate one invariant only, cases may exist[15] in which a complete set of invariants can be generated, yielding a solution similar to Lacina's but free from secularity.

REFERENCES

(1) H. Goldstein, Classical Mechanics, Addison-Wesley 1953

(2) H. Corben and P. Stehle, Classical Mechanics, 2nd Ed., John Wiley 1960

(3) Ref.(2), Section 74

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

(5) D. Stern, On Kruskal's Perturbation Method (preceding this article). The notation in that work differs from the one used here in that n there denotes the total number of variables.

(6) J. Lacina, Czech. J. of Physics B-19, 130 (1969)

(7) J. Lacina, Annals of Physics 51, 381 (1969)

(8) Sophus Lie, Theorie der Transformationgruppen I, Teubner, Leipzig, 1888

(9) G. Hori, Publications of the Astronomical Society of Japan 18, 287 (1966)

(10) A. Deprit, Celestial Mechanics 1, 12 (1969)

(11) P. Musen, J. of the Astronautical Sciences 12, 129 (1965)

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

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

(14) G. Contopoulos, Astronomical Journal 68, 1 (1963)

(15) Ref. (5), eq. (24) and text following.