X-641-70-363
PREPRINT

NASA TM X- 65366

# DIRECT NEAR-IDENTITY CANONICAL TRANSFORMATIONS


DAVID P. STERN



FACILITY FORM 602

N70-42579 (ACCESSION NUMBER) (THRU)

10 (PAGES) 1 (CODE)

TMX-65366 (NASA CR OR TMX OR AD NUMBER) 19 (CATEGORY)


OCTOBER 1970



GSFC

GODDARD SPACE FLIGHT CENTER
GREENBELT, MARYLAND

PRESENTED AT THE 12th ASTRODYNAMICS CONFERENCE
GODDARD SPACE FLIGHT CENTER
OCTOBER 26-27, 1970

DIRECT NEAR-IDENTITY CANONICAL TRANSFORMATIONS

David P. Stern
Laboratory for Space Physics
Goddard Space Flight Center

Presented at the 12th Astrodynamics Conference
GSFC, 26-27 October 1970

Abstract

Let $\underline{y} = (\underline{p}, \underline{q})$ be a canonical set of variables undergoing a near-identity transformation that involves a small parameter

$$\underline{y} \rightarrow \underline{z} = \underline{y} + \sum \varepsilon^k \underline{\zeta}^{(k)}(\underline{y})$$

Then the transformation will be canonical provided $\underline{\zeta}^{(k)}$ has the form

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

where $\bar{\nabla}$ is a gradient operator in "conjugate phase space" $\bar{\underline{y}} = (\underline{q}, -\underline{p})$, $\chi^{(k)}$ is arbitrary and may be regarded as the k-th order of a generating function and $\underline{f}^{(k)}$ is a given vector depending on lower orders of $\underline{\zeta}$ and on their derivatives. Various choices of $\underline{f}$ exist (all differing by some conjugate gradient) and some of these may be related to the conventional form or to Lie's form of canonical transformations. Properties and uses of this representation will be discussed.

Note: Equations are numbered according to the slide on which they appear in the presentation.

I would like to present here a somewhat new approach to the old subject of canonical transformations.

I am sure everyone here is quite familiar with such transformations. If we start with a system of canonical variables

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

having N degrees of freedom (that is, $\underline{y}$ has 2N components) and seek a transformation to a new system

$$\underline{z} = (\underline{P}, \underline{Q}) \tag{1-1}$$

then the usual way of expressing this is by means of a <u>generating function</u> $\sigma$ depending on N "new" and N "old" variables. There exist many forms for this, but we shall choose the form

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

This gives the equations of transformation as

$$\begin{aligned} Q_i &= \partial\sigma/\partial P_i \\ p_i &= \partial\sigma/\partial q_i \end{aligned} \tag{1-2}$$

Now among all the various canonical transformations, a particularly important type consists of the <u>near-identity transformations</u>, depending on a <u>small parameter $\varepsilon \ll 1$</u> in a way which reduces them to identity transformations $\underline{y} = \underline{z}$ as $\varepsilon \to 0$ . A generating function for such a transformation may have the form

$$\sigma(\underline{P}, \underline{q}) = \sum P_i q_i + \sum \varepsilon^k \sigma^{(k)}(\underline{P}, \underline{q}) \tag{1-3}$$

from which

$$Q_i = q_i + \sum \varepsilon^k \partial\sigma^{(k)}/\partial P_i \tag{1-4}$$

$$p_i = P_i + \sum \varepsilon^k \partial\sigma^{(k)}/\partial q_i \tag{1-5}$$

This form is the basis of the familiar Poincaré-Von Zeipel method: if we transform without time dependence, the new Hamiltonian $H^*(\underline{P}, \underline{Q})$ equals the old one $H(\underline{p}, \underline{q})$ , and by substituting the above formulas we get

$$H\left(P_i + \sum \varepsilon^k (\partial \sigma^{(k)}/\partial q_i),\ q_i\right) =$$

$$= H^*\left(P_i,\ q_i + \sum \varepsilon^k (\partial \sigma^{(k)}/\partial P_i)\right) \qquad (1\text{-}6)$$

This relation involves only the 2N variables $(\underline{P}, \underline{q})$ and by expanding, demanding that $H^*$ does not contain the angle variable and that $\sigma^{(k)}$ has no secular parts, solutions to problems in perturbed periodic motion may be worked out order by order.

Unfortunately, this transformation from the original variables to more desirable ones is expressed by functions of both old and new variables, and one still must untangle this relationship order by order.

It would be much better if we had a <u>direct</u> transformation, of the form

$$\underline{z} = \underline{y} + \sum \varepsilon^k \underline{\zeta}^{(k)}(\underline{y}) \qquad (2\text{-}1)$$

or, for many uses

$$\underline{y} = \underline{z} + \sum \varepsilon^k \underline{\eta}^{(k)}(\underline{z}) \qquad (2\text{-}2)$$

in which all new variables are explicitely given in terms of old ones, or vice versa.

One way of getting such a transformation is by <u>Lie's method</u>. About 80 years ago Lie showed that if $L_W$ is a Poisson-bracket operator

$$L_W(f) \equiv \left[ f, W \right] \qquad (2\text{-}3)$$

then the transformation

$$\underline{z} = \exp(\varepsilon L_W) * \underline{y} \qquad (2\text{-}4)$$

(asterisk means operation) is a near-identity canonical one. The exponential is here defined by its Taylor expansion and may be reduced to a series like the one in equation (2-1) (or else, interchanging $\underline{z}$ and $\underline{y}$ gives a series like 2-2 ). If furthermore $W$ can be expanded in powers of $\varepsilon$

$$W = \sum \varepsilon^k W^{(k)}(\underline{y}) \qquad (2\text{-}5)$$

then each order of the expansion will introduce the corresponding order of $W$, and will in addition contain assorted expressions involving lower orders.

This may then be substituted into the Poincare-Von Zeipel method, that is

$$H(\underline{y}) \quad = \quad H^*(\underline{y} + \sum \varepsilon^k \underline{\zeta}^{(k)})$$

and again separated into a set of expressions, one for each order. Just as ordinarily each order contains an undetermined function $\sigma^{(k)}$ which one adjusts to eliminate the angle variable and secular terms, so here one does it with $W^{(k)}$ . This is essentially the method advocated by Hori in 1966 and later developed by Deprit.

Sometimes, of course, one prefers to deal directly with the form given in equations (2-1) and (2-2)

$$\underline{z} \quad = \quad \underline{y} \quad + \quad \sum \varepsilon^k \underline{\zeta}^{(k)}(\underline{y}) \tag{2-1}$$

$$\underline{y} \quad = \quad \underline{z} \quad + \quad \sum \varepsilon^k \underline{\eta}^{(k)}(\underline{z}) \tag{2-2}$$

For instance, the Bogoliubov-Krylov method (or its mirror image, used by Kruskal) leads to such a transformation; if $\underline{y}$ is a canonical set, one may then ask under what conditions is the transformation also canonical.

I want to devote the rest of the talk to this question : as you will see, it leads to a representation which includes both Lie transforms and the conventional generating function. We will deal with the form (2-1), but everything can be equally well done with the inverse series (2-2).

Note first of all that eq. (2-1) deals with the vectors $\underline{y}$ and $\underline{z}$ without splitting them into canonical momenta and coordinates. We will try to preserve this unity throughout the calculation, and it can be done — but for a price: one must introduce "conjugate" vectors.

Let us define them: if

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

then "y-conjugate" is

$$\bar{\underline{y}} \quad = \quad (\underline{q},\ -\underline{p}) \tag{3-1}$$

That is, $\bar{y}$ is formed by rearranging the components of $\underline{y}$ so that every $p_i$ is replaced by its conjugate and every $q_i$ by minus its conjugate.

With this notation, Hamilton's equations simply become

$$dy_i/dt \quad = \quad - \quad \partial H/\partial \bar{y}_i \tag{3-2}$$

or, in vector form

$$dy/dt = - \bar{\nabla} H \qquad (3\text{-}3)$$

Poisson brackets involve only one summation

$$[a, b] = \sum \left( \frac{\partial a}{\partial q_i} \frac{\partial b}{\partial p_i} - \frac{\partial a}{\partial p_i} \frac{\partial b}{\partial q_i} \right)$$

$$= \sum \frac{\partial a}{\partial \bar{y}_i} \frac{\partial b}{\partial y_i} = \bar{\nabla} a \cdot \nabla b \qquad (3\text{-}4)$$

In particular

$$[a, y_i] = \partial a / \partial \bar{y}_i \qquad (3\text{-}5)$$

and

$$[\bar{y}_i, y_j] = \delta_{ij} \qquad (3\text{-}6)$$

which is the condition for a set $\underline{y}$ to be canonical.

<u>Now back to the main question.</u> If $\underline{y}$ is canonical and $\underline{z}$ differs from it only by small terms, the condition for $\underline{z}$ to be canonical is

$$[z_i, z_j] - [y_i, y_j] = 0 \qquad (4\text{-}1)$$

Substituting

$$\left[ y_i + \sum{}^k \zeta_i^{(k)},\; y_j + \sum \varepsilon^m \zeta_j^{(m)} \right] - [y_i, y_j] = 0 \qquad (4\text{-}2)$$

This separates into a series of relations, one for each order. The zeroth order cancels, of course. The $O(\varepsilon)$ part gives

$$[\zeta_i^{(1)}, y_j] - [\zeta_j^{(1)}, y_i] = 0 \qquad (4\text{-}3)$$

or, by equation (3-5)

$$\partial \zeta_i^{(1)} / \partial \bar{y}_j - \partial \zeta_j^{(1)} / \partial \bar{y}_i = 0 \qquad (4\text{-}4)$$

This looks very much like the curl dyadic of $\zeta^{(1)}$ in $\bar{\underline{y}}$ space, and in fact we shall designate it as such and write

$$(\bar{\nabla} \times \underline{\zeta}^{(1)})_{ij} = 0 \qquad (4\text{-}5)$$

One can easily guess the solution: if the curl of a vector vanishes, it is the gradient of some scalar, so

$$\underline{\zeta}^{(1)} = \bar{\nabla} \chi^{(1)} \qquad (4\text{-}6)$$

with $\chi^{(1)}$ arbitrary.

<u>In general</u> one finds, for the k-th order

$$(\bar{\nabla} \times \underline{\zeta}^{(k)})_{ij} = - \sum_{m=1}^{k-1} \left[ \zeta_i^{(m)}, \ \zeta_j^{(k-m)} \right] \qquad (5\text{-}1)$$

Since all terms on the right are of orders lower than the k-th, this may serve as a recursion relation. Let us denote by $\underline{f}^{(k)}$ any solution for $(k)$ in this equation — a solution which, in general, will depend on lower orders of $\zeta$ and on their derivatives. Then one can always add to this solution an arbitrary conjugate gradient $\bar{\nabla} \chi^{(k)}$, since the curl operation anyway wipes it out. Thus the general solution will have the form

$$\underline{\zeta}^{(k)} = \bar{\nabla} \chi^{(k)} + \underline{f}^{(k)}(\underline{\zeta}) \qquad (5\text{-}2)$$

with $\underline{f}^{(k)}$ any particular solution. For instance, one possible form for the next two orders turns out to be

$$\underline{\zeta}^{(2)} = \bar{\nabla} \chi^{(2)} + \tfrac{1}{2} \underline{\zeta}^{(1)} \cdot \nabla \underline{\zeta}^{(1)} \qquad (5\text{-}3)$$

$$\underline{\zeta}^{(3)} = \bar{\nabla} \chi^{(3)} + \underline{\zeta}^{(2)} \cdot \nabla \underline{\zeta}^{(1)} \qquad (5\text{-}4)$$

The vectors $\underline{f}^{(k)}$ are not unique — given any one set of such vectors, we can always form another one by adding to it expressions having the form of conjugate gradients. However, if we have <u>any one such set</u> available, all near-identity transformations can be characterised by it.

For lower orders, one can construct solutions by vector manipulation, but this quickly becomes difficult. However, since any direct canonical transformation must have the form (5-2), <u>any</u> method for obtaining such a transformation must lead to a choice of $\underline{f}^{(k)}$ .

Fortunately, there exist at least two such methods.

First of all, there's the Lie transform. If one expands it and derives a formula for $\underline{\zeta}^{(k)}$, one finds

$$\underline{\zeta}^{(k)} = - \bar{\nabla} W^{(k)} + \text{(expression involving lower orders)}$$

Obviously, $-W^{(k)}$ is our $\chi^{(k)}$ and the remaining terms give one choice for $\underline{f}^{(k)}$.

Alternatively, one can untangle the transformation based on $\sigma(\underline{P}, \underline{q})$ and express it as a direct transformation. The result is

$$\underline{\zeta}^{(k)} = - \bar{\nabla} \sigma^{(k)} + \text{(expression involving lower orders)}$$

Again, this gives a choice for $\underline{f}^{(k)}$ and, as it turns out, this form is simpler to derive and shorter than the other one. Any of these forms can be substituted in the Poincaré-Von Zeipel method, giving

$$H(\underline{y}) = H^{*}\left( \underline{y} + \sum \varepsilon^{k} \left[ \bar{\nabla}\chi^{(k)} + \underline{f}^{(k)} \right] \right)$$

Depending on our choice of $\underline{f}^{(k)}$, we get from this the various orders either of the "conventional" or the "Lie - type" generating function that transforms our variables to a set in which the Hamiltonian does not contain the angle variable.

(incidentally, if we want to use the inverse expansion 2-2, we must substitute in $H$, not in $H^{*}$)

The method can also be extended to situations with a slow explicit dependence on time; the constant transformed action variable then goes by the name of adiabatic invariant. Since in such cases we need the "conventional" generating function for deriving the new Hamiltonian, we must use the direct transformation based on $\sigma$ and not, say, that derived from Lie transforms.

In conclusion, let me say that I suspect that this approach to near-identity canonical transformations offers more convenience than others. For those interested, I have copies of an article on the subject that just appeared in J. Math. Physics.

What else it means, I don't know, though it seems that the concept of conjugate coordinates has something behind it — perhaps some canonical formalism involving complex numbers. The use of the conjugate gradient and curl also suggest an interesting geometry in phase space, and I hope all this will be investigated in due course.

(end)

Afterthoughts on the Application of Direct Canonical Transformations

(not presented at the meeting)

The new forms of canonical transformations — those described here and Lie's — provide alternative possibilities for new perturbation schemes replacing the traditional method of Poincaré and Von Zeipel. While all these methods lead to equivalent results, I believe — for reasons outlined below — that the traditional method still offers the most convenient approach.

Any of these methods derives a near-identity canonical transformation, characterized in each order by the appropriate order of some "generating function" ( $\sigma^{(k)}$ of (1-3), $W^{(k)}$ of (2-5), $\chi^{(k)}$ of (5-2), etc.), and the calculation usually centers around the derivation, order by order, of this function and of the transformed Hamiltonian $H^*$ . If the transformation is a direct one, the task ends here; in the traditional method, however, an inversion must still be performed to untangle the relationship between old and new variables and to bring it to a "direct" form.

Suppose we used a "direct" transformation

$$\underline{z} = \underline{y} + \sum \varepsilon^k \underline{\zeta}^{(k)}(\underline{y}) \tag{A-1}$$

with

$$\underline{\zeta}^{(k)} = \bar{\nabla}\chi^{(k)} + \underline{f}^{(k)} \tag{A-2}$$

Then, as on page 7, the basic relation is

$$\begin{aligned} H(\underline{y}) &= H^*(\underline{z}) \\ &= H^*\left(\underline{y} + \sum \varepsilon^k \left[\bar{\nabla}\chi^{(k)} + \underline{f}^{(k)}\right]\right) \end{aligned} \tag{A-3}$$

and this must now be expanded order by order as in the ordinary Poincaré-Von Zeipel method. Different choices of $\underline{f}^{(k)}$ lead to different perturbation methods: for instance, choosing as in (5-3)

$$\underline{f}^{(2)} = \tfrac{1}{2}\, \underline{\zeta}^{(1)} \cdot \nabla \underline{\zeta}^{(1)} \tag{A-4}$$

( $\underline{f}^{(1)}$ vanishes) gives something equivalent to the Lie transform method. On the other hand, let the notation

$$\underline{\zeta}^{(k)} = - \bar{\nabla}\sigma^{(k)}(\underline{p}, \underline{q}) + \underline{\rho}^{(k)} \qquad \text{(A-5)}$$

denote the particular direct transformation corresponding to the "conventional" transformation generated by $\sigma(\underline{P}, q)$ of (1-3), after it is untangled (this is the transformation mentioned as the 2nd alternative on top of page 7). With this choice

$$\underline{f}^{(2)} \equiv \underline{\rho}^{(2)} = \sum \left( \frac{\partial\sigma^{(1)}}{\partial q_i} \frac{\partial}{\partial p_i} \right) \underline{\zeta}^{(1)} \qquad \text{(A-6)}$$

Comparison shows that (A-6) has only half as many differentiations as (A-4) and therefore ought to be easier to use. Similar simplifications exist in higher orders and for that reason, of all the direct transformations of form (A-2), that one which is given in (A-5) appears to be the most convenient one for use in a perturbation theory.

If instead of all this we used the traditional Poincaré-Von Zeipel method, we would have expanded (1-6) and derived from it $\sigma^{(k)}(\underline{P}, q)$ and $H^*$ order by order. Having derived $\sigma^{(k)}(\underline{P}, q)$, we would then invert the relation and derive the corresponding direct transformation, simply by writing $\underline{p}$ in place of $\underline{P}$ everywhere in the generating function and substituting the functions $\sigma^{(k)}(\underline{p}, \underline{q})$ thus obtained in (A-5).

This procedure derives the <u>same</u> function (within the freedom allowed by the method) as does (A-3) with the choice (A-5) for the direct transformation -- which, as has been pointed out, is probably the most convenient choice for (A-3). The difference is that in the traditional method, $\sigma^{(k)}$ is first derived to all orders and <u>then</u> $\rho^{(k)}$ is formed and the direct transformation (A-5) is obtained. By contrast, in the direct form of the method (A-3), the vectors $\rho^{(k)}$ (taking the role of $\underline{f}^{(k)}$) are already introduced in the expansion of the Hamiltonian and have to be carried along throughout the <u>entire</u> calculation.

One may thus conclude that the traditional method is more convenient, since it introduces $\underline{f}^{(k)}$ only in the end. Of course, the difference only appears for $k \geq 2$, because $\underline{f}^{(1)}$ vanishes.

The formulas for deriving $\rho^{(k)}$ are listed on the following page.

Note on the use of (A-5)

The general formula for $\underline{\rho}^{(k)}$ is (see "Direct Canonical Transformations", to appear in J. Math. Phys.)

$$\underline{\zeta}^{(k)} = - \sum_{m=1}^{k-1} S^{(m)} * \bar{\nabla}\sigma^{(k-m)} \qquad \text{(A-7)}$$

where $S^{(m)}$ are expansion operators (the asterisk denotes their operation), with the first few being

$$S^{(1)} = \underline{\pi}^{(1)} \cdot \nabla$$

$$S^{(2)} = \underline{\pi}^{(2)} \cdot \nabla + \tfrac{1}{2}\, \underline{\pi}^{(1)} \underline{\pi}^{(1)} : \nabla\nabla \qquad \text{(A-8)}$$

$$S^{(3)} = \underline{\pi}^{(3)} \cdot \nabla + \underline{\pi}^{(1)} \underline{\pi}^{(2)} : \nabla\nabla + (1/6)\, \underline{\pi}^{(1)} \underline{\pi}^{(1)} \underline{\pi}^{(1)} \therefore \nabla\nabla\nabla$$

and so forth, and where $\underline{\pi}^{(s)}$ is just the momentum-like part of the vector $\underline{\zeta}^{(s)}$, that is

$$\underline{\pi}^{(s)} = ( \zeta_1^{(s)}, \ldots \zeta_N^{(s)}, 0, \ldots 0 ) \qquad \text{(A-9)}$$

For instance, for the order following (A-6) one gets ( $\underline{\zeta}^{(1)} = - \bar{\nabla}\sigma^{(1)}$)

$$\underline{\zeta}^{(3)} = \underline{\pi}^{(2)} \cdot \nabla \underline{\zeta}^{(1)} + \tfrac{1}{2}\, \underline{\pi}^{(1)} \underline{\pi}^{(1)} : \nabla\nabla \underline{\zeta}^{(1)} - ( \underline{\pi}^{(1)} \cdot \nabla ) ( \bar{\nabla} \sigma^{(2)} ) \qquad \text{(A-10)}$$

This can be expressed directly in terms of the $\sigma^{(s)}$ alone, but the result is somewhat more cumbersome than the preceding formula.