Computer Algebra Implementation of Lie Transforms for Hamiltonian Systems: Application to the Nonlinear Stability of $L_4$

We apply the method of Lie Transforms, a perturbation method for differential equations, to a general class of Hamiltonian systems using computer algebra. In doing so, we develop explicit formulas (written in terms of coefficients in the original system $H$) for transforming the system into Birkhoff normal form. Using a theorem of Arnold, we develop explicit nonlinear stability criteria solely in terms of $H$ for systems where the linear stability analysis is inconclusive. After providing two examples, we apply our results to the nonlinear stability of $L_4$, a triangular point in the circular restricted three body problem. At $L_4$, Arnold's theorem must be used since a Lyapunov function cannot be found.

Introduction

This work concerns Lie transforms, a method for obtaining approximate solutions to systems of differential equations. We apply the method to a general class of two degree of freedom Hamiltonian systems, viz., two coupled nonlinear oscillators with nonresonant frequencies. For systems in this class, we use Lie transforms to approximately reduce the system to an equivalent simpler system which is immediately solvable, i.e., a system with ignorable coordinates.

As an application of our results, we determine the nonlinear stability of the triangular points in the circular restricted three body problem. In doing so we corroborate a computation recently performed by Meyer and Schmidt [16]. Their computation was based on their own computer algebra program written in PL/I, whereas the present work is based on readily available utilities written in MACSYMA [19]. Moreover, while their computation was specifically performed for the problem at the triangular point $L_4$, the present work applies to a problem with arbitrary (symbolic) coefficients.

We begin by introducing the reader to Lie transforms. Then we show how the method may be applied to a particular class of problems, and finally we specialize the results to some examples, including the problem at $L_4$.

Lie transforms

In this section we summarize the method of Lie transforms (see [8], [12], [15], [17]). This work is concerned with Hamiltonian systems, i.e., systems which are derivable from a single scalar function $H$, the Hamiltonian:

$$\frac{dx_m}{dt} = \partial H / \partial y_m, \quad \frac{dy_m}{dt} = -\partial H / \partial x_m,$$

where $x_m$ and $y_m$ are the dependent variables of the problem, $m = 1, \ldots, N$, where $N$ is called the number of degrees of freedom. The method of Lie transforms generates a near-identity transformation from $(x_m, y_m)$ to $(X_m, Y_m)$ variables,

$$x_m = X_m + \text{quadratic terms in } (X_k, Y_k) + \text{cubic terms in } (X_k, Y_k) + \ldots,$$

$$y_m = Y_m + \text{quadratic terms in } (X_k, Y_k) + \text{cubic terms in } (X_k, Y_k) + \ldots,$$

which is canonical, i.e., which preserves the Hamiltonian form of the equations:

$$\frac{dX_m}{dt} = \partial K / \partial Y_m, \quad \frac{dY_m}{dt} = -\partial K / \partial X_m,$$

where $K = K(x_m, Y_m) = H(x_m, y_m)$ is the Hamiltonian in the new variables (called the Hamiltonian after Goldstein [11]).

The near-identity transformation is generated by first introducing a scaling parameter $\varepsilon$ into the problem. Expanding $H$ in a power series about the origin (assumed to be an equilibrium position) yields

$$H = H_0(x_m, y_m) + \varepsilon H_1(x_m, y_m) + \varepsilon^2 H_2(x_m, y_m) + \cdots,$$

in which the $H_i$ are polynomials, and $\varepsilon$ is a small parameter.
where \( H_n(x_m, y_m) \) is a polynomial of degree \( n + 2 \). Then the near-identity transformation is generated by the associated Hamiltonian system
\[
\frac{dx_m}{de} = \partial W / \partial y_m, \quad \frac{dy_m}{de} = -\partial W / \partial x_m,
\]
in which \( e \) plays the role of time. The transformation evolves in \( e \), starting with the initial conditions
\[
e = 0, \quad x_m = X_m, \quad y_m = Y_m.
\]
The Hamiltonian \( W \) of equations (5), called the generating function, is also expanded in a power series in \( e \):
\[
W = W_1 + e W_2 + e^2 W_3 + \ldots,
\]
where \( W_n \) is a polynomial of degree \( n + 2 \). The point of this generating scheme is that the resulting transformation is canonical for any choice of the \( W_n \)'s (see [8], [15]). The actual choice of these functions depends upon the problem at hand, but the main idea is to pick them so that the new Hamiltonian \( K \) is as simple as possible. We note that the parameter \( e \) in this paper corresponds to \( -e \) in [15] and [19].

The transformation is generated by expanding the variables \((x_m, y_m)\) in Taylor series in \( e \) and using the generating equations (5)–(7) to evaluate the coefficients,
\[
x_m = x_m|_{e=0} + \frac{d x_m}{d e} \bigg|_{e=0} e + \frac{d^2 x_m}{d e^2} \bigg|_{e=0} e^2 + \ldots,
\]
\[
y_m = y_m|_{e=0} + \frac{d y_m}{d e} \bigg|_{e=0} e + \frac{d^2 y_m}{d e^2} \bigg|_{e=0} e^2 + \ldots.
\]

The transformation is thus found to be given by
\[
x_m = X_m + \frac{\partial W_1}{\partial Y_m} e + \left[ \frac{\partial W_2}{\partial Y_m} + \left( \frac{\partial W_1}{\partial X_m}, W_1 \right) \right] e^2 + \ldots
\]
and similarly,
\[
y_m = Y_m + \frac{\partial W_1}{\partial X_m} e + \left[ \frac{\partial W_2}{\partial X_m} + \left( \frac{\partial W_1}{\partial X_m}, W_1 \right) \right] e^2 + \ldots.
\]

In order to obtain the transformed Hamiltonian \( K \) (cf. (3)), the transformation (13), (14) is substituted into a power series expansion for the original Hamiltonian \( H \):
\[
K(X_m, Y_m) = H(x_m, y_m) = H_0(x_m, y_m) + e H_1(x_m, y_m) + e^2 H_2(x_m, y_m) + \ldots,
\]
\[
H_0(x_m, y_m) = H_0 \left( X_m + \frac{\partial W_1}{\partial Y_m} e + \ldots, Y_m - \frac{\partial W_1}{\partial X_m} e - \ldots \right) = H_0|_{e=0} + \bigg| \frac{d H_0}{d e} \bigg|_{e=0} e + \frac{d^2 H_0}{d e^2} \bigg|_{e=0} e^2 + \ldots,
\]
\[
H_0|_{e=0} = H_0(X_m, Y_m),
\]
\[
\frac{d H_0}{d e} \bigg|_{e=0} = \sum_j \frac{\partial H_0}{\partial x_j} \frac{d x_j}{d e} + \frac{\partial H_0}{\partial y_j} \frac{d y_j}{d e} \bigg|_{e=0} = \sum_j \frac{\partial H_0}{\partial x_j} \frac{\partial W_1}{\partial y_j} + \frac{\partial H_0}{\partial y_j} \frac{\partial W_1}{\partial x_j} \bigg|_{e=0} = \sum_j \partial H_0 \partial x_j \partial W_1 \partial Y_j - \partial H_0 \partial Y_j \partial x_j = \{ H_0, W_1 \}.
\]

This equation, which represents the expansion of \( H_0 \) under the near-identity transformation (13), (14), also holds for any of the \( H_n \)'s, and in fact is valid for any function \( f(x_m, y_m) \). Substitution of (19) and the corresponding equations on the other \( H_n(x_m, y_m) \) into (15) gives, after some simplification:
\[
K(X_m, Y_m) = K_0(X_m, Y_m) + K_1(X_m, Y_m) e + K_2(X_m, Y_m) e^2 + \ldots,
\]
where

\[ K_0 = H_0, \quad K_1 = H_1 + \{ H_0, W_1 \}, \quad K_2 = H_2 + \frac{1}{2} \{ H_0, W_2 \} + \frac{1}{2} \{ K_1, W_1 \} + \frac{1}{2} \{ H_1, W_1 \}, \]

(21), (22), (23)

\[ K_3 = H_3 + \frac{1}{3} \{ H_0, W_3 \} + \frac{1}{3} \{ K_1, W_2 \} + \frac{1}{3} \{ K_2, W_1 \} + \frac{1}{3} \{ H_1, W_2 \} + \frac{2}{3} \{ H_2, W_1 \} + \]

+ \frac{1}{3} \{ (H_3, W_1), W_1 \}, \]

(24)

\[ K_4 = H_4 + \frac{1}{4} \{ H_0, W_4 \} + \frac{1}{4} \{ K_1, W_3 \} + \frac{1}{4} \{ K_2, W_2 \} + \frac{1}{4} \{ K_3, W_1 \} + \]

+ \frac{1}{4} \{ (H_2, W_2), W_1 \} + \frac{3}{4} \{ (H_3, W_1), W_1 \} + \frac{1}{4} \{ (H_1, W_2), W_1 \} + \frac{1}{4} \{ (H_2, W_1), W_1 \} + \frac{1}{4} \{ (H_2, W_1), W_1 \}. \]

(25)

In equations (21)–(25), the \( H_n \) and \( W_n \) are taken as functions of the variables \( X_m, Y_m \).

So we see that the method of Lie transforms is nothing more than the introduction of the generating equations (5)–(7) into Taylor series expansions for the variables \( (x_m, y_m) \) and \( H \). However, the transformation equations (e.g. (21)–(25)) can be generated much more efficiently than by the foregoing expansion method. There are several schemes for doing so (including the original method of DEFRIT [8] based on the “Lie triangle” and a method of DRAGT and FINS [10] based on infinite products rather than infinite series), but we prefer the following method (see [15]), which is easily implemented on MACSYMA ([13], [14], [19]).

Define the operators \( L_n \) and \( S_n \) as follows:

\[ L_n = (, W_n), \]

(26)

\[ S_0 = \text{Id} \ (\text{the identity operator}), \quad S_n = \frac{1}{n} \sum_{m=0}^{n-1} L_{n-m} S_m, \quad n = 1, 2, 3, \ldots \]  

(27.1), (27.2)

Then the near-identity transformation from \( (x_m, y_m) \) to \( (X_m, Y_m) \) variables is given by

\[ x_m = [S_0 + \varepsilon S_1 + \varepsilon^2 S_2 + \ldots] X_m, \quad y_m = [S_0 + \varepsilon S_1 + \varepsilon^2 S_2 + \ldots] Y_m \]

(28.1), (28.2)

and the \( n \)-th term \( K_n \) of the Hamiltonian is given by the expression

\[ K_n = H_n + \frac{1}{n} \{ H_0, W_n \} + \frac{1}{n} \sum_{m=1}^{n-1} [L_{n-m} K_m + mS_{n-m} H_m], \quad n = 2, 3, 4, \ldots \]  

(29)

where the cases \( n = 0, 1 \) are given by equations (21), (22).

Coupled oscillators

In this work we shall apply the method of Lie transforms to two degree of freedom Hamiltonian systems in which \( H_0 \) has the special form:

\[ H_0 = \frac{1}{2} (p_1^2 + \omega_1^2 y_1^2) - \frac{1}{2} (p_2^2 + \omega_2^2 y_2^2), \]

(30)

where \( q_m \) and \( p_m \) are variables representing the displacement and momentum of oscillator \( m \). For \( \varepsilon = 0 \), the equations of motion corresponding to such a Hamiltonian become

\[ \dot{q}_m = p_m \quad \text{and} \quad \dot{p}_m = -\omega_m^2 q_m, \quad \text{or} \quad \ddot{q}_m + \omega_m^2 q_m = 0. \]

(31)

Thus when \( \varepsilon = 0 \), the system has eigenvalues \( \pm i \omega_m \), where \( i = \sqrt{-1} \), and we change variables to eigencoordinates \( (x_m, y_m) \),

\[ q_m = \frac{x_m}{\omega_m} + \frac{y_m}{2}, \quad p_m = \frac{\omega_m y_m}{2} + i x_m \]

(32)

for which the equations of motion (31) and Hamiltonian (30) take the form

\[ \dot{x}_m = \omega_m x_m \quad \text{and} \quad \dot{y}_m = -\omega_m y_m, \]

(33)

\[ H_0 = \omega_m x_m y_m. \]

(34)

In these coordinates, each \( H_n \) becomes a polynomial of degree \( n + 2 \) in the four variables \( x_1, y_1, x_2, y_2 \). For example, there are 20 cubic monomials which form a basis for \( H_4 \):

\[ H_4 = \text{linear combination of } \{ x_1^2, x_1^2 x_2, x_1^2 y_1, x_1^2 y_2, x_1 x_2 y_1, x_1 x_2 y_2, x_1 y_1^2, x_1 y_1 y_2, x_1 y_2^2, x_2^2 y_1, x_2^2 y_2, x_2 y_1^2, x_2 y_1 y_2, y_1^2, y_1 y_2, y_2^2 \}. \]

(35)

The numbers of basis monomials for \( H_2, H_3, \) and \( H_4 \) are:

<table>
<thead>
<tr>
<th>Term</th>
<th>Degree</th>
<th>No. of basis monomials</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_1 )</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>( H_2 )</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>( H_3 )</td>
<td>5</td>
<td>56</td>
</tr>
<tr>
<td>( H_4 )</td>
<td>6</td>
<td>84</td>
</tr>
</tbody>
</table>
We now come to the question of how to choose the generating functions $W_n$ so as to best simplify the Kamiltonians $K_n$. At the $n$-th step of the method, $K_n$ is given by equation (29),

$$K_n = \frac{1}{n} \{H_0, W_n\} + \text{terms which are already known}. \tag{36}$$

Now with $H_0$ in the simplified form (34),

$$\{H_0, W_n\} = \frac{\delta H_0 \delta W_n}{\delta Y_1 \delta Y_1} - \frac{\delta H_0 \delta W_n}{\delta X_1 \delta X_1} - \frac{\delta H_0 \delta W_n}{\delta Y_2 \delta X_2} =$$

$$= i\omega_1 \left[ Y_1 \frac{\partial W_n}{\partial Y_1} - X_1 \frac{\partial W_n}{\partial X_1} \right] - i\omega_2 \left[ Y_2 \frac{\partial W_n}{\partial Y_2} - X_2 \frac{\partial W_n}{\partial X_2} \right] \tag{37}$$

we want to choose $W_n$ so that this linear partial differential operator on $W_n$ cancels as many terms as possible in (36). Each term to be cancelled will be of the form

$$AX^l Y_1^r Y_2^s \tag{38}$$

where $A$ is a constant. In view of the linearity of (37), we choose $W_n$ to be a sum of terms, one for each term (38) to be cancelled, of the form

$$W_n = BX^l Y_1^r Y_2^s \tag{39}$$

where $B$ is an undetermined constant. Then

$$\frac{1}{n} \{H_0, W_n\} = \frac{1}{n} (i\omega_1 (l-j) - i\omega_2 (s-r)) BX^l Y_1^r Y_2^s \tag{40}$$

leading to the choice

$$B = \frac{iAn}{\omega_1 (l-j) - \omega_2 (s-r)}, \quad n = j + l + r + s - 2. \tag{41}$$

Note that this scheme fails if the denominator of (41) vanishes. Assuming that the frequencies $\omega_1$ and $\omega_2$ are incommensurable (nonresonant), the denominator will vanish only if both

$$l = j \quad \text{and} \quad s = r. \tag{42}$$

Thus we cannot remove terms of the form

$$(X_1 Y_1)^l (X_2 Y_2)^r. \tag{43}$$

This means that we can always reduce every such (nonresonant) problem to the form:

$$K_0 = H_0 = i\omega_1 (X_1 Y_1) - i\omega_2 (X_2 Y_2), \quad K_1 = 0, \tag{44}, (45)$$

$$K_2 = K_{300}(X_1 Y_1)^2 + K_{111}(X_1 Y_1)(X_2 Y_2) + K_{002}(X_2 Y_2)^2, \quad K_3 = 0, \tag{46}, (47)$$

$$K_4 = K_{300}(X_1 Y_1)^3 + K_{211}(X_1 Y_1)^2(X_2 Y_2) + K_{112}(X_1 Y_1)(X_2 Y_2)^2 + K_{002}(X_2 Y_2)^3. \tag{48}$$

That is, every such nonresonant two degree of freedom problem can, to $O(4)$, be reduced to only 7 coefficients. Note that in this case the resulting Hamiltonian is a function only of the "action" variables,

$$I_1 = iX_1 Y_1 \quad \text{and} \quad I_2 = iX_2 Y_2 \tag{49}$$

and hence both coordinates are ignorable and the system is immediately solvable to $O(4)$. Such a system is said to be in Birkhoff normal form ([5], p. 85).

By inspection of (41), the foregoing scheme fails at special resonant values of $\omega_1$ and $\omega_2$. In solving for $W_n$, resonant terms occur for integer values of $k_1$ and $k_2$ such that

$$k_1 \omega_1 + k_2 \omega_2 = 0, \quad |k_1| + |k_2| \leq n + 2. \tag{50}$$

In such cases additional non-removable terms occur. We shall not consider such resonant cases in this work.

**Computer algebra**

The computation just described turns out to involve vast quantities of algebra. We used the computer algebra system MACSYMA ([18]) in order to do the computation more accurately and more efficiently than by hand. For example, the key formulas (12), (26), (27), (29) can be represented in MACSYMA via the following lines of code ([7], [19]):

```
POISSON(F,G):= 
SUM(DIFF(F,X(1))DIFF(G,Y(1)) - DIFF(F,Y(1))DIFF(G,X(1)),I,1,N)
L(IF):=POISSON(F,W[I])
S(IF):=(IF I=0 THEN F ELSE SUM(L(I-M,S(M,F)),M,O,I-1))
K[I]:=(H[I]+POISSON(H[0],W[I])
+SUM(L(I-M,K[M])+M*S(I-M,H[M]),M,1,I-1))
```
In order to efficiently compute \( W_n \) by the formulas (39), (41), we use the MACSYMA tool called pattern matching. A rule named WSOLVE is defined as follows:

\[
\text{LET}(X^1 \rightarrow Y^1 \cdot L \cdot X^2 \rightarrow R \cdot Y^2 \rightarrow S, \\
X^1 \cdot Y^1 \cdot L \cdot X^2 \rightarrow R \cdot Y^2 \rightarrow S, \alpha_n I \cdot N(W^1 \cdot (L - J) \cdot W^2 \cdot (S - R)), \text{WSOLVE})$
\]

That is, replace the term \( X^i Y^i X^j Y^j \) by \( \frac{\text{in}X^i Y^i X^j Y^j}{\alpha_n(l - j) - \alpha_n(s - r)} \). When WSOLVE is applied to the “terms which are already known” on the right hand side of equation (36), the correct expression for \( W_n \) is automatically generated. Note that this rule is not applied to non-removable terms of the form (43).

One could hope to simply apply these formulas to the problem at hand, and to thereby automatically obtain the transformed Kamiltonian. Unfortunately, the size of the \( O(4) \) computation is too large to proceed directly; MACSYMA on a Symbolics 3670 runs out of space. E.g., from (25) we see that the computation of \( K_4 \) involves the evaluation of the quantity \( \{ \{ H_1, W_1 \}, W_1 \}, W_1 \}. The innermost Poisson bracket involves 20 terms for \( H_1 \) and 20 terms for \( W_1 \), i.e. 400 pairs which can be collected together into 35 terms (since there are 35 fourth degree basis monomials). These then need to be combined with the 20 terms of \( W_1 \) in order to evaluate the second Poisson bracket, i.e. 700 pairs which come together into 56 terms. Next the third Poisson bracket combines the previous result with the 20 terms of \( W_1 \) to require the computation of 1120 pairs, which may be collected together into 84 terms.

In order to complete the computation, we broke it up into pieces, each of which was sufficiently small so as not to cause MACSYMA to encounter space problems. We shall refer to our strategy for treating such large computations as the method of telescoping compositions. As an example of this strategy, we once again consider the computation of the triple Poisson bracket \( \{ \{ H_1, W_1 \}, W_1 \}, W_1 \}. We first compute \( \{ H_1, W_1 \} \) and store the resulting 35 coefficients \( A_{jtr} \) in a disk file. Next, instead of computing \( \{ \{ H_1, W_1 \}, W_1 \} \), we compute instead \( \{ A, W_1 \} \), where \( A \) is a dummy polynomial with symbolic coefficients \( A_{jtr} \). Although we are eventually interested in identifying these coefficients with those we have stored in a disk file, we save that step for later. We store the resulting 56 coefficients \( B_{jtr} \) of \( \{ A, W_1 \} \) in a disk file. Next we compute \( \{ B, W_1 \} \), where now \( B \) is a dummy polynomial with symbolic coefficients \( B_{jtr} \). This results in 84 coefficients which are known in terms of the \( B_{jtr} \) coefficients. The latter are stored in a file and are known in terms of the \( A_{jtr} \) coefficients, which are also stored in a disk file. At this point the computation of \( \{ \{ H_1, W_1 \}, W_1 \}, W_1 \} \) is complete, although it still remains to plug the values of the \( A_{jtr} \) and \( B_{jtr} \) coefficients into the final result. — For a complete listing of the programs, see [7].

Results

The results of this work take the form of expressions for the transformed Kamiltonian \( K \) in terms of the original Hamiltonian \( H \). If we express \( H \) in \( x_m, y_m \) eigencoordinates defined by equations (32), then \( H_0 \) takes the canonical form (34), and the polynomials \( H_n \) of (4) can be written as

\[
H_n = \sum_{m} H_{jtr} x_m^j y_m^r z_m^s,
\]

where the \( H_{jtr} \) are given constants. Then the coefficients \( K_{jtr} \) in \( K_4 \) in equation (46) are given by:

\[
K_{2000} = H_{2000} + i \left( \frac{1}{\omega_2} H_{1001} H_{1110} + \frac{3}{\omega_1} \left( H_{1200} H_{2100} + H_{0300} H_{3000} \right) - \frac{1}{2 \omega_1 + \omega_2} H_{0210} H_{2001} + \frac{1}{2 \omega_1 - \omega_2} H_{0201} H_{2001} \right).
\]

\[
K_{1111} = H_{1111} + i \left( \frac{2}{\omega_2} \left( H_{1011} H_{1200} + H_{0111} H_{2100} \right) - \frac{4}{\omega_1 + 2 \omega_2} H_{0102} H_{1002} - \frac{2}{\omega_2} \left( H_{1011} H_{0021} + H_{1110} H_{0012} \right) + \frac{4}{\omega_2 + 2 \omega_1} H_{0210} H_{0021} - \frac{4}{2 \omega_2 - \omega_1} H_{0102} H_{0020} + \frac{4}{2 \omega_1 - \omega_2} H_{0210} H_{0020} \right).
\]

\[
K_{0022} = H_{0022} + i \left( \frac{1}{\omega_1} H_{0111} H_{1011} - \frac{3}{\omega_2} \left( H_{0003} H_{0030} + H_{0012} H_{0021} \right) + \frac{1}{2 \omega_1 + \omega_2} H_{0120} H_{0022} - \frac{1}{2 \omega_1 - \omega_2} H_{0102} H_{0020} \right).
\]

The comparable coefficients in \( K_4 \) in (48) were also found, but cannot be displayed here because they are too long. E.g., the ASCII files for \( K_{2000} \) and \( K_{0003} \) contain 164K characters, while those for \( K_{1111} \) and \( K_{0022} \) contain 468K. These expressions simplify greatly, however, in the special case in which \( H_1 \) and \( H_2 \) are identically zero. Since this
special case occurs in frequently in sample problems, we give the associated coefficients of \( K_4 \) here:

\[
K_{3000} = H_{3000} \quad \frac{iH_{3010}H_{3010}}{a_2 - 3a_1} - \frac{iH_{3101}H_{3110}}{a_2 - a_1} - \frac{iH_{3210}H_{3210}}{a_2 + a_1} - \frac{iH_{3010}H_{3001} + 4iH_{3100}H_{4000} + 4iH_{3100}H_{3100}}{a_1} \quad (54)
\]

\[
K_{2211} = H_{2211} \quad \frac{9iH_{3010}H_{3010}}{a_2 - 3a_1} - \frac{i(3H_{3101} + 2H_{3110})H_{3110}}{a_2 - a_1} + \frac{9iH_{3100}H_{3001}}{a_2 + 3a_1} + \frac{i(3H_{3110} - 2H_{3110})H_{3100}}{a_2 + a_1} - \frac{2iH_{3202}H_{3202}}{a_2 - a_1} - \frac{2iH_{3212}H_{3212}}{a_2 - a_1} + \frac{2iH_{3202}H_{3002}}{a_2 + a_1} - \frac{2iH_{3212}H_{3102}}{a_2 + a_1} - \frac{2iH_{3121}H_{3121}}{a_2 - a_1} + \frac{3iH_{3211}H_{3210}}{a_1} + \frac{3iH_{3300}H_{3300}}{a_1} \quad (55)
\]

\[
K_{1122} = H_{1122} \quad \frac{9iH_{3010}H_{3010}}{3a_2 - a_1} - \frac{i(3H_{3012} + 2H_{3101})H_{3101}}{3a_2 - a_1} + \frac{9iH_{3100}H_{3000}}{3a_2 + a_1} - \frac{i(3H_{3012} - 2H_{3101})H_{3100}}{3a_2 - a_1} - \frac{2iH_{3202}H_{3202}}{a_2 - a_1} - \frac{2iH_{3212}H_{3212}}{a_2 - a_1} + \frac{2iH_{3202}H_{3002}}{a_2 + a_1} - \frac{2iH_{3212}H_{3102}}{a_2 + a_1} - \frac{2iH_{3121}H_{3121}}{a_2 - a_1} - \frac{3iH_{3013}H_{3100}}{a_1} - \frac{3iH_{3031}H_{3100}}{a_1} \quad (56)
\]

\[
K_{0033} = H_{0033} \quad \frac{iH_{3013}H_{3030}}{3a_2 - a_1} + \frac{iH_{3103}H_{3030}}{3a_2 + a_1} - \frac{iH_{3013}H_{3021} + iH_{3013}H_{1012}}{a_2 - a_1} + \frac{iH_{3013}H_{1012}}{a_2 + a_1} - \frac{4iH_{3006}H_{0006}}{a_2} - \frac{4iH_{0006}H_{0002}}{a_2} \quad (57)
\]

Arnold's theorem

We are interested in applying the previous results to the determination of the stability of the equilibrium at the origin in a system of two nonlinear coupled oscillators in which \( H_0 \) has the form (30). Note that the linearized Hamiltonian differential equations (1) corresponding to \( H = H_0 \) have purely imaginary eigenvalues, and thus are inconclusive regarding stability. Moreover, because of the minus sign in (30), \( H_0 \) is not positive definite, and Lyapunov's direct method [7] cannot be used to determine stability.

For such cases, stability may be determined by appealing to a theorem of ARNOLD [4], which has been restated and reproved by MEYER and SCHMIDT [16]. The theorem, based on the existence of invariant tori in KAM theory [3], gives sufficient conditions for stability in nonresonant systems, in terms of the transformed Hamiltonian \( K (J_1, J_2) \) which has been put in Birkhoff normal form, cf. (49). The terms \( K_n \) of equations (44)–(48) are thought of as functions of \( J_1 \) and \( J_2 \), \( K_n (J_1, J_2) \). The theorem involves quantities \( D_n \) defined by

\[
D_n = K_n(a_2, a_1). \quad (58)
\]

From (44)–(48), the first two non-identically zero \( D_n \)'s are \( D_2 \) and \( D_4 \):

\[
D_2 = -(K_{3000}a_2^2 + K_{1111}a_1a_2 + K_{2222}a_1^2), \quad D_4 = i(1K_{3000}a_2^2 + K_{2211}a_1a_2^2 + K_{1122}a_1^2a_2 + K_{0033}a_1^3). \quad (59)
\]

Arnold's theorem states that the origin is stable for those parameter values for which \( D_2 \neq 0 \). In the case that \( D_2 = 0 \), stability is assured if \( D_4 \neq 0 \), and so on. I.e., the origin is stable if \( D_{2n} \neq 0 \) for some \( n \).

Using the expressions (51)–(57) for the coefficients \( K_{jkr} \), expressions for \( D_2 \) and \( D_4 \) (the latter in the special case that \( H_1 = H_3 = 0 \)) may be obtained:

\[
D_2 = -(\omega_2^2H_{3000} + \omega_1\omega_2H_{1111} + \omega_1^2H_{0022}) + \frac{3\omega_2^2}{a_2} (H_{1102}H_{1110} - \omega_1H_{1002}H_{1110} + 2\omega_1H_{1110}H_{0022} + H_{1102}H_{1102}) - 2\omega_2H_{1101}:H_{0020} + H_{1101}H_{1200}) + \frac{3\omega_2^2}{a_2} (H_{1102}H_{0020} + H_{1102}H_{1200}) + \frac{4\omega_2 + \omega_1}{2a_2} \omega_1H_{1002}H_{1022} + \frac{4\omega_2 + \omega_1}{2a_2} \omega_1H_{1002}H_{1020} - \frac{4\omega_2 + \omega_1}{2a_2} \omega_2H_{1010}H_{0210} - \frac{4\omega_2 + \omega_1}{2a_2} \omega_2H_{0001}H_{0210}. \quad (61)
\]
\[ D_4 = i(\omega_2 H_{2300} + \omega_2^* H_{2111} + \omega_2 \omega_4 H_{1123} + \omega_2^* H_{0032}) - 2 \omega_2 \omega_4 H_{1103} H_{2110} + H_{0112} H_{2011} + H_{0203} H_{2010} + H_{1011} H_{2101} + H_{1112} H_{2110} + H_{2020} H_{2002} - \\
+ H_{0113} H_{2110} + H_{2101} H_{0121} - 3 \omega_2^2 (H_{2000} H_{0211} + H_{1300} H_{0021}) - 3 \omega_2^* (H_{0012} H_{1120} + H_{0030} H_{1122}) - \\
\frac{4 \omega_1^2}{\omega_3} (H_{0004} H_{0040} + H_{0013} H_{0021}) - \frac{4 \omega_2^2}{\omega_3} (H_{0006} H_{0040} + H_{1300} H_{3100}) - \\
\frac{\omega_4^2 H_{1201} H_{2110}(w_2 + 3 \omega_4)}{\omega_2 + \omega_4} - \frac{\omega_4^2 H_{1210} H_{2110}(w_2 - 3 \omega_4)}{w_2 - \omega_4} - \\
\frac{\omega_4^2 H_{0112} H_{0121}(w_1 + 3 \omega_2)}{w_1 + \omega_2} - \frac{\omega_4^2 H_{0310} H_{0210}(w_2 - 3 \omega_4)}{w_2 - 3 \omega_4} - \\
\frac{\omega_4^2 H_{0103} H_{1020}(w_1 + 9 \omega_4)}{w_1 + 3 \omega_2} - \frac{\omega_4^2 H_{0130} H_{0300}(w_2 - 9 \omega_4)}{w_2 - 3 \omega_4}. \tag{62} \]

(assumes \( H_1 = H_3 = 0 \)).

The expression for \( D_4 \) in the general case is too long to be included here, but is available on our computer for numerical evaluation.

**Example 1**

We consider a variation of the Henon-Heiles Hamiltonian where the linear oscillators are not at low-order resonance and are of different signs:

\[ H = \frac{1}{2} (p_1^2 + \omega q_1^2) - \frac{1}{2} (p_2^2 + \omega q_2^2) + \frac{1}{2} q_3^2. \tag{63} \]

Using the transformation to eigencoordinates given by equation (32), \( H \) becomes

\[ H = i \omega x_1 y_1 - i x_2 y_2 - \frac{1}{3} x_2^3 + \frac{i}{24} y_2^3 + \frac{i}{20} x_2 y_2 + \frac{1}{4} y_2^2 x_2 - \frac{i}{2} x_2 y_2 - \frac{i}{4} y_2^2 x_2 + \]

\[ + \frac{i}{\omega} x_2 y_2 - \frac{1}{20} x_2 y_2, \quad \omega > 0. \tag{64} \]

Then using equations (51)–(63), we find the \( K_2 \) coefficients to be

\[ K_{2000} = \frac{3 - 8 \omega^2}{4 \omega^2(4 \omega^2 - 1)}, \quad K_{1111} = \frac{4 \omega^2 + 1}{\omega(4 \omega^2 - 1)}, \quad K_{0022} = - \frac{5}{12}. \tag{65}, (66), (67) \]

Using equation (61), we find \( D_z \) to be

\[ D_z = \frac{2 \omega^2 - 53 \omega^4 + 12 \omega^2 - 9}{12 \omega^2(4 \omega^2 - 1)}. \tag{68} \]

We find that \( D_z = 0 \) only for \( \omega = \omega_c \approx 1.5752978 \). In order to determine the stability of the origin for \( \omega = \omega_c \), we must consider the \( D_4 \) condition. Because \( H \) contains only cubic nonlinear terms and because each cubic coefficient is simple, we are able to find the expression for \( D_4 \) algebraically. The coefficients for \( K_4 \) turn out to be

\[ K_{3300} = \frac{i(1024 \omega^6 + 780 \omega^8 + 1632 \omega^4 + 596 \omega^2 - 51)}{48 \omega^2(2 \omega - 1)^2(2 \omega + 1)^2}, \quad K_{2211} = \frac{i(3840 \omega^{10} - 288 \omega^8 + 16 \omega^6 - 340 \omega^4 + 159 \omega^2 - 6)}{4 \omega^2(2 \omega^2 - 1)^2(2 \omega - 1)^2}, \tag{69}, (70) \]

\[ K_{1122} = \frac{i(4 \omega^2 - 1) (320 \omega^6 - 480 \omega^4 + 360 \omega^2 - 161 \omega^2 + 6)}{12 \omega^2(2 \omega - 1)^2(2 \omega + 1)^2}, \quad K_{0033} = - \frac{235 i}{432}, \tag{71}, (72) \]

and \( D_4 \) becomes

\[ D_4 = \frac{(15040 \omega^{18} - 72400 \omega^{14} + 113172 \omega^{12} - 77350 \omega^{10} + 14491 \omega^8 - 10188 \omega^6 - 3096 \omega^4 + 5175 \omega^2 - 459)}{432 \omega^2(2 \omega - 1)^2(2 \omega + 1)^2}. \tag{73} \]

So, \( \omega = \omega_c, D_z = -0.19180289 \neq 0 \).

Thus by Arnold’s theorem, the origin is nonlinearly stable. We note that this result does not apply to a small set of resonant values of \( \omega \) which correspond to vanishing denominators in the algorithm (41). From equation (50) with \( n = 2 \), we find the following resonant values of \( \omega \):

\[ \omega = (\frac{1}{2}, \frac{1}{3}, 1, 2, 3). \]

**Example 2**

This second example involves a spinning mass-spring system, which contains no odd powered terms in the Hamiltonian. Consider 4 identical springs, each attached at one end to the outer rim of a wheel of unit radius separated by 90°. The other end of each spring is attached to a unit mass which is free to move about its equilibrium position at the center, see Fig. 1. Let the \( Q_1 = Q_2 \) axes rotate with the wheel with angular velocity \( \omega > 0 \) relative to an inertial frame. Each spring is unstretched when the mass is at the origin.
The potential energy $V_i$ for each spring under a deflection $\delta_i$ is taken to be

$$V_i = \frac{1}{2} (\delta_i^2 + \mu \delta_i^4)$$

(74)

where the linear spring constant has been taken equal to $\frac{1}{2}$, and $\mu$ is a nonlinear spring constant. Then this system has the Hamiltonian

$$H = \frac{1}{2} (P_1^2 + P_2^2) + \omega (P_1 Q_2 - P_2 Q_1) + V_1 + V_2 + V_3 + V_4$$

(75)

where $P_i$ are momenta. Then, upon taking the Taylor series of $H$ about the origin, $H$ becomes [7]

$$H = \frac{1}{2} (P_1^2 + P_2^2) + \omega (P_1 Q_2 - P_2 Q_1) + \frac{1}{2} (Q_1^2 + Q_2^2) + \frac{1}{4} [(4\mu + 1) Q_1^4 - 8Q_1^2 Q_2^2 + (4\mu + 1) Q_2^4] -$$

$$- \frac{1}{12} [Q_1^6 + 4(\mu - 1) Q_1^4 Q_2^2 + 4(\mu - 1) Q_2^4 Q_1^2 + Q_2^6] + \ldots = H_0 + H_3 + H_4 + \ldots$$

(76)

Using the linear differential equations corresponding to $H_0$, we find the characteristic equation to be

$$\lambda^2 + 2\lambda (1 + \omega^2) + (\omega^2 - 1)^2 = 0$$

(77)

which has eigenvalues $\lambda = \left(\pm i (1 + \omega), \pm i (1 - \omega)\right)$. From this we conclude that the equilibrium at the origin is elliptic for $\omega \neq 1$, i.e., comprised of two oscillators with frequencies $1 - \omega$ and $1 + \omega$ in the first approximation. Then using a canonical eigenvector transformation from $(Q_m, P_m)$ to $(x_m, y_m)$ gives [7]

$$H_0 = i(1 - \omega) x_1 y_1 - i(1 + \omega) x_2 y_2$$

(78)

which is in the proper form for our analysis. After similarly transforming $H_2$ and $H_4$, we use equations (51)–(57) to find that

$$K_{2200} = K_{0022} = \frac{1 - 12\mu}{32}, \quad K_{1111} = \frac{1 - 12\mu}{8}$$

(79)

$$K_{3300} = \frac{i(576\mu^2 - 32\mu + 20) \omega^2 - (864\mu^2 - 48\mu + 30) \omega + 272\mu^2 - 56\mu - 15)}{1024(\omega - 1)(2\omega - 1)}$$

(80)

$$K_{2211} = \frac{3i(1440\mu^2 - 48\mu + 58) \omega^2 - (720\mu^2 - 24\mu + 29) \omega - 48\mu^2 - 120\mu - 75)}{1024\omega(2\omega - 1)}$$

(81)

$$K_{1122} = \frac{3i(1440\mu^2 - 48\mu + 58) \omega^2 + (720\mu^2 - 24\mu + 29) \omega - 48\mu^2 - 120\mu - 75)}{1024\omega(2\omega + 1)}$$

(82)

$$K_{0020} = \frac{i(576\mu^2 - 32\mu + 20) \omega^2 + (864\mu^2 - 48\mu + 30) \omega + 272\mu^2 - 56\mu - 15)}{1024(\omega + 1)(2\omega + 1)}$$

(83)

We find $D_2$ and $D_4$ using eqs. (61)–(62) to be

$$D_2 = \frac{(12\mu - 1)(3\omega^2 - 1)}{16}$$

(84)

$$D_4 = \left[\frac{(9792\mu^2 - 352\mu + 388) \omega^4 - (17744\mu^2 - 264\mu + 1213) \omega^8 + (9104\mu^2 + 232\mu + 657) \omega^4}{- (1392\mu^2 + 216\mu + 207) \omega^4 - 144\mu^2 - 36\mu - 225)} \right] \frac{1}{512\omega(\omega^2 - 1)(4\omega^2 - 1)}$$

(85)

Then $D_2 = 0$ for $\mu = \frac{1}{3}$ and $\omega^2 = \frac{3}{5}$, which are two lines in the $\mu - \omega$ parameter plane. When $D_2 = 0$ we must check the $D_4$ condition. Consider the line $\mu = \frac{1}{3}$. The value of $D_4$ on this line is

$$D_4 \left(\mu = \frac{1}{3}\right) = \frac{60\omega^4 - 191\omega^4 + 104\omega^4 - 33\omega^4 - 36}{72\omega(\omega^2 - 1)(4\omega^2 - 1)}$$

(86)

which is zero only for $\omega = \omega_c \approx 1.6241875 \ldots$. Now consider the line $\omega^2 = \frac{3}{5}$. $D_4$ on this line becomes

$$D_4 \left(\omega^2 = \frac{3}{5}\right) = \frac{\sqrt{3}}{288}(336\mu^2 + 1064\mu + 661)$$

(87)

which is zero only for $\mu = \mu_{\omega^2 = \frac{3}{5}} = -\frac{1}{3} (133 \pm \sqrt{133}) \approx (-0.84870, -2.31796)$. We now apply the stability theorem. First, note that we consider $\omega > 0$ and that for $\omega = 1$ the origin is not elliptic so that our analysis does not apply there. From equation (50) with $n = 2$, we must also exclude $\omega = \{\frac{1}{3}, \frac{1}{5}, 2, 3\}$ from the analysis. Applying the $D_4$ condition, we find that the origin is stable everywhere in the $\mu - \omega$ parameter plane except possibly along the two lines $\mu = \frac{3}{4}$ and $\omega^2 = \frac{3}{5}$. On these lines the $D_4$ condition must be used. From (50) with $n = 4$, we must now exclude $\omega = \{\frac{1}{2}, \frac{2}{3}, \frac{1}{4}, \frac{2}{5}, \frac{3}{5}, \frac{5}{7}\}$ when $\mu = \frac{3}{10}$. Elsewhere on $\mu = \frac{3}{10}$, the origin is stable provided $\omega \neq \omega_c$; on $\omega^2 = \frac{3}{5}$, the origin is stable provided $\mu \neq \mu_{\omega^2 = \frac{3}{5}}$. For the three points where $D_2 = D_4 = 0$, the $D_4$ condition must be used to prove nonlinear stability.

We note that for $\omega < 1$, stability of the origin can be independently proved by Lyapunov's direct method [7].
Application to the problem of three bodies

The circular restricted three body problem is well-known to exhibit five equilibria in a rotating barycentric coordinate system [20]. $L_1$, $L_2$, and $L_3$ represent equilibrium positions of the third body, in which all three bodies are collinear. All three of these are unstable for all values of the mass ratio parameter $\mu$. $L_4$ and $L_5$ represent equilibria where all three bodies sit at the vertices of an equilateral triangle. For values of $\mu > \mu_1 \simeq 0.0385208$, both these equilibria are unstable. For $\mu < \mu_1$, ALFRIEND [1, 2] showed that the triangular points are unstable when $\mu = \mu_2$ and $\mu_3$, special mass ratios which cause the linearized frequencies to be in the ratio of $1:2:1:3$, respectively. For other values of $\mu < \mu_3$, stability of $L_4$ and $L_5$ can be obtained by using Arnold's theorem. This was first done by DEPRIT and DEPRIT-BARTHOLOME [9], who calculated $\mu$ by hand. The value they obtained,

$$D_2 = -\frac{36 - 5410\omega_2^2 + 6440\omega_1^2\omega_2^4}{1 - 4\omega_1^2\omega_2^2} \left( 4 - 25\omega_2^2 \right)$$

is non-zero for all values of $\mu$ except for $\mu = \mu_4 \simeq 0.0109136$. For $\mu = \mu_4$, Arnold's theorem requires the quantity $D_4$ be found. This computation was performed by MEYER and SCHMIDT, who found $D_4 \simeq -66.6$. The non-zero value of $D_4$ implies stability, by Arnold's theorem.

In what follows we shall apply the results obtained in this paper to confirm the previous computations of DEPRIT and DEPRIT-BARTHOLOME [9] and MEYER and SCHMIDT [16].

The Hamiltonian for the circular restricted three-body problem about the equilibrium $L_4$ is:

$$H = \frac{1}{2}(P_1^2 + P_2^2) + P_1Q_2 - P_2Q_1 - \frac{(1 - 2\mu)}{2}Q_1 - \sqrt{3}Q_2 - \left( \mu_1 + 1 - \mu, \mu_2 \right),$$

where $\omega_1^2 = Q_1^2 + Q_2^2 - Q_1 + \sqrt{3}Q_2 + 1$, $\omega_2^2 = Q_1^2 + Q_2^2 + Q_1 + \sqrt{3}Q_2 + 1$. Expanding in a Taylor series about the origin, $H$ becomes $H_{\mu}$, where $H_{\mu}$ contains terms of order $n + 2$ and $H_5$ is given by:

$$H_5 = \frac{1}{2}(P_1^2 + P_2^2) + P_1Q_2 - P_2Q_1 + \frac{1}{8}Q_1^2 - Q_2^2 - \frac{3\sqrt{3}}{4}(1 - 2\mu)Q_1Q_2.$$  

Then using the linearized differential equations corresponding to $H_{\mu}$, the characteristic equation for the system is found to be:

$$\lambda^4 + \lambda^2 + \frac{27}{16}(1 - \gamma^2) = 0 \quad \text{where} \quad \gamma = 1 - 2\mu.$$  

The eigenvalues $\lambda$ have positive real parts for $\mu > \mu_1 = \frac{1}{2}\left(1 - \frac{\sqrt{69}}{9}\right)$ implying the system is linearly unstable. For $\mu < \mu_1$, the system is critically stable having eigenvalues $\pm i\omega_1$ and $\pm i\omega_2$ where:

$$0 < \omega_2 < \frac{\sqrt{2}}{2} < \omega_1 < 1, \quad \omega_1^2 + \omega_2^2 = 1, \quad \omega_1^2\omega_2^2 = \frac{27}{16}(1 - \gamma^2).$$

Using a canonical linear transformation (see [6]), $H_{\mu}(q_m, p_m)$ is transformed into $H_{\mu}(q_m, p_m)$ which is of the form (30). Then following equations (30)-(34) we introduce the variables $(x_m, y_m)$ and find the following components of $K_2$:

$$K_{2222} = -\frac{\omega_1^2(124\omega_1^4 - 996\omega_1^2 + 58)}{144(2\omega_1^2 - 1)^2(\omega_2^2 - 4\omega_2)}, \quad K_{1} = \omega_1^2(64\omega_1^2 - 64\omega_1^2 - 43)$$

and the first stability condition becomes:

$$D_2 = -\frac{-6440\omega_1^8 + 1288\omega_1^6 - 1185\omega_1^4 + 5410\omega_1^2 - 36}{8(2\omega_1^2 - 1)^2(\omega_2 - 2\omega_1)(\omega_2 - 2\omega_1)(2\omega_2 - \omega_1)(2\omega_2 + \omega_1)}$$

which is equivalent to the expression (88) found by DEPRIT and DEPRIT-BARTHOLOME [9]. Then, on $0 < \mu < \mu_1$, $D_2 = 0$ only for:

$$\mu = \mu_c = \frac{3}{483} - \frac{\sqrt{2(199945 + 326)}}{6483} \simeq 0.0109136.$$  

At this value of $\mu = \mu_c$, the components of $K_4$ become:

$$K_{2222} \simeq 0.219259187i + 6.52E-37, \quad K_{2} \simeq -7.79324843i + 3.74E - 35,$$

and $D_4$ becomes:

$$D_4 \simeq -66.6 - 4.27E-36i.$$  

(99)
The very small real part of each $K_{\mu_0}$ and imaginary part of $D_4$ results from taking only a finite number of digits (40 in fact) in the numerical approximation. Because the real part is so much larger than the error term, the approximation $D_3 \approx -66.6$ is accurate and $D_4 \approx 0$. Hence, at $\mu = \mu_0$, the Hamiltonian system is stable. These values for the coefficients of $K_1$ and $D_4$ agree with those obtained by Meyer and Schmidt [16].

Acknowledgement: The authors would like to thank Dr. Dieter Armbruster for his helpful advice with MACSYMA. This work was partially supported by grants from the National Science Foundation, AFOSR, and ARO through the Mathematical Sciences Institute at Cornell.

References


Received May 2, 1988

Address: Dr. Richard H. Rand, Professor; Dr. Vincent T. Coppola, Department of Theoretical and Applied Mechanics College of Engineering, Cornell University, Thurston Hall, Ithaca, NY 14853–1503, USA