Unrestricted Harmonic Balance
A General Method to Evaluate Periodic Structures in Time and/or Space of Arbitrary Stability in Non-linear Differential Equation Systems

V. Investigation of the Truncation Errors in the Case of an Exactly Solvable Non-linear System

Friedrich Franz Seelig
Institut für Physikalische und Theoretische Chemie der Universität Tübingen, Lehrstuhl für Theoretische Chemie, Tübingen

Z. Naturforsch. 38a, 729–735 (1983); received March 28, 1983

The method of Unrestricted Harmonic Balance (UHB) is a generally applicable procedure to
determine the Fourier coefficients and constants of motion (frequency, velocity) of non-linear periodic phenomena in time and space and was outlined and applied to various problems in chemistry in preceding papers. Here the truncation error is investigated by comparing the UHB results for various truncation harmonics N with the rigorous analytical solution for the Kepler problem which is one of the rare examples of an exactly solvable non-linear problem.

1. Introduction

In the preceding 4 papers of this series the method of Unrestricted Harmonic Balance (UHB) was developed and demonstrated for non-linear oscillations [1], applied to stiff ordinary differential equations in enzyme catalysis [2], used to get standing- and running waves in partial differential equations of chemical reaction-diffusion systems [3], and extended to transcendental functions as they occur in chemical kinetics, if the exponential dependence of the rate constants on the temperature is considered [4]. The main features of UHB base on the fact that periodic phenomena can rigorously be expressed by an (infinite) Fourier series, e.g. for oscillations of some state variable \( x(t) \) with period \( T \) and fundamental frequency \( \omega = 2 \pi/T \) as

\[
x(t) = \tilde{x} + \sum_{j=1}^{\infty} x_{cj} \cos(j \omega t) + \sum_{j=1}^{\infty} x_{sj} \sin(j \omega t).
\]  

(1)

If the non-linearities consist of products of state variables each product, say \( p(t) \), is again of the same form (1) and behaves thus like a pseudo-linear variable, but now \( \tilde{p} \) and the \( \{p_{cj}\}, \{p_{sj}\} \) are complicated non-linear functions of \( \tilde{x}, \{x_{cj}\}, \{x_{sj}\}, \tilde{y}, \{y_{cj}\}, \{y_{sj}\} \) of the factors \( x(t) \) and \( y(t) \).

In practical applications the Fourier series have to be truncated at some highest harmonic, say \( N \), so that each state variable \( x(t) \) is characterized by an \((2N - 1)\)-dimensional vector with components \( \tilde{x}, \{x_{cj}\}, \{x_{sj}\} \) and the product formation \( p(t) = x(t)y(t) \) can be expressed by a particular vector product yielding a \((2N + 1)\)-vector from two \((2N + 1)\)-vectors as factors. Nevertheless this method is "unrestricted", because \( N \) is flexible and can be increased in subsequent passes. The procedure was programmed in a subroutine and can be easily applied every time it occurs in the formulas. In the case of transcendental functions the function is expanded in a Taylor or McLaurin series as usual, but each power of \( x \) is gained from the preceding power by application of the multiplication subroutine. The final result of the manipulations is in every case that a system of non-linear algebraic equations rather than differential equations is gained which can be solved iteratively by conventional methods, e.g. the method of Powell [5].

If the (exact) Fourier series is truncated at harmonic \( N \), a truncation error results. If two series of this truncated kind are multiplied, the product has its highest harmonic of order \( 2N \), but the method needs a further truncation to \( N \) again. So – especially if repeated multiplications occur – the truncation errors accumulate, in principle at least. Intuitively the decrease of amplitudes from \( j = 1 \) to \( j = N \) was used to control \( N \) for a given accuracy, but criticism pointed just to that deficiency.

In order to get some insight into the mechanism of error proliferation a rigorously solvable non-

Reprint requests to Prof. F. F. Seelig, Institut für Physikalische und Theoretische Chemie der Universität Tübingen, Auf der Morgenstelle 8, 7400 Tübingen.

0340-4811 / 83 / 0700-0729 S 01.3 0/0. – Please order a reprint rather than making your own copy.
linear sample system is needed whose exact solution can be compared to the results of UHB.

2. The Kepler Problem as a Rigorously Solvable Example

Since generally non-linear differential equations cannot be solved analytically and a mere comparison with other approximate methods, e.g. the Runge-Kutta-Merson simulation technique does not seem appropriate, finding a rigorously solvable example turned out to be a difficult task. But since the UHB-method claims to be quite general the domain of chemical kinetics was abandoned for this purpose.

The Kepler problem of the description of the movement of the earth or some other planet around the attracting sun turned out to be the wanted example and is given in the textbook on theoretical physics by Landau and Lifschitz [6].

There are different ways for expressing the equation of motion, e.g. balance of attraction (gravitation) and repulsion (centrifugation) forces leading to second order differential equations with corresponding initial conditions. Here another formulation utilizing certain conservation laws appeared to be more appropriate.

The elliptic revolution of the earth around the sun is shown in Figure 1. $a$ and $b$ are long and short axis of the ellipsis, $e$ is the numerical excentricity. The sun with mass $M$ is located in one focus of the ellipsis, the earth with mass $m$ has the coordinates $x(t)$ and $y(t)$ and components of velocity $\dot{x}(t)$ and $\dot{y}(t)$, the distance sun-earth is $r = \sqrt{x^2 + y^2}$. Then the total energy $E$ as sum of kinetic and potential (gravitational) energy of the earth is a constant of motion and given by

$$ E = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2) - \gamma (Mm/r) ,$$

where $\gamma$ is the gravitational constant.

Another constant of motion is the $z$-component of angular momentum

$$ Q_z = m (x \dot{y} - y \dot{x}) .$$

Defining $e$ as energy per mass unit of earth and $q$ as angular momentum per mass unit we get

$$ e = E/m = \frac{1}{2} (\dot{x}^2 + \dot{y}^2) - \gamma M/r$$

and

$$ q = Q_z/m = x \dot{y} - y \dot{x} .$$

Besides, $q$ is double the area spread over by the radius vector per unit time which is constant, too (Kepler’s second law). The problem can be simplified without loss of generality, if dimensionless variables are introduced. Bearing in mind that attracting states with negative energy are considered we choose

$$ \xi = -2e x/\gamma M ,$$

$$ \eta = -2e y/\gamma M ,$$

$$ \tau = (-2e)^{3/2} t/\gamma M$$

and need now only one parameter, namely

$$ x = (-2e)^{1/2} q/\gamma M .$$

Fig. 1a. Ellipsis of the Kepler problem of the revolution of a planet around the sun. Assumed excentricity $e = 0.43589$ corresponding to $x = 0.9$. The points are equidistant in time and reflect the varying velocity.

Fig. 1b. Dimensionless coordinates $\xi$ and $\eta$ vs. $\tau$ for one period showing the non-linearity.
With
\[ \dot{x} = \frac{d\xi}{dt} = \frac{dx}{dt} + \frac{d\xi}{d\tau} \frac{d\tau}{dt} = \frac{d\xi}{d\tau} \frac{d\tau}{dt} = (-2e)^{1/2} \frac{d\xi}{d\tau} \]
(4) becomes
\[ -\frac{1}{2} \left[ \left( \frac{d\xi}{d\tau} \right)^2 + \left( \frac{d\eta}{d\tau} \right)^2 \right] = -\frac{1}{\sqrt{\xi^2 + \eta^2}} \]
(11)
and (5) transforms to
\[ \frac{\dot{\xi}}{d\tau} = \frac{\xi}{d\tau} - \frac{\eta}{d\tau} \cdot \frac{d\xi}{d\tau} \]
(12)
The solution is tried in a parametric representation, according to Landau-Lifschitz:
\[ \xi = a(e + \cos \varphi) , \]
(13)
\[ \eta = b \sin \varphi , \]
(14)
\[ \tau = c(\varphi + \xi \sin \varphi) \]
(15)
with
\[ b = a \sqrt{1 - e^2} \]
(16)
since \( \sin \varphi = \frac{\eta}{b} \) and \( \cos \varphi = \frac{\xi - a e}{a^2} \), and \( \frac{(\xi - a e)^2}{a^2} + \frac{\eta^2}{b^2} = 1 \) describes an ellipsis. \( \varphi \) goes from 0 to 2\( \pi \) as \( \tau \) goes from 0 to \( T = 2\pi/\omega \) so that from (15) \( c \) is seen to be \( 1/\omega \).

Insertion of (13), (14), (15) into (12) and regarding
\[ \frac{d\xi}{d\tau} = \frac{d\xi}{d\varphi} \frac{d\varphi}{d\tau} = \frac{d\xi}{d\varphi} / \frac{d\varphi}{d\tau} \]
(17)
yields
\[ \zeta = \frac{a(e + \cos \varphi) b \cos \varphi}{c(1 + e \cos \varphi)} + \frac{b \sin \varphi a \sin \varphi}{c(1 + e \cos \varphi)} = \frac{ab}{c} . \]
(18)
The same procedure on (11) leads to
\[ -\frac{1}{2} \left[ \left( \frac{d\xi}{d\tau} \right)^2 + \left( \frac{d\eta}{d\tau} \right)^2 \right] = -\frac{1}{\sqrt{\xi^2 + \eta^2}} \]
\[ = \frac{1}{2} \left( \frac{\sin^2 \varphi}{c^2(1 + e \cos \varphi)^2} + \frac{\cos^2 \varphi}{c^2(1 + e \cos \varphi)^2} \right) \]
\[ = \frac{a^2 + 2e \cos \varphi + \cos^2 \varphi + (1 - e^2) \sin^2 \varphi}{2c^2(1 + e \cos \varphi)^2} \]
\[ = \frac{1}{2} \frac{a^2(1 - e \cos \varphi) - 1}{a(1 + e \cos \varphi)} \]
(19)
or
\[ -\frac{1}{2} \left( 1 + e \cos \varphi \right) - \frac{1}{2} \frac{a^2}{c^2(1 - e \cos \varphi) + 1} = 0. \]
(20)
This to be true for arbitrary \( \varphi \) needs
\[ -\frac{1}{2} + \frac{1}{2} \frac{a^2}{c^2} = 0 \]
(21)
and
\[ \frac{1}{2} - \frac{1}{2} \frac{a^2}{c^2} + \frac{1}{a} = 0 . \]
(22)
So for given constants of motion, \( e \) and \( q \), i.e. given \( \zeta \)
\[ a = 1 \]
(23)
and
\[ \omega = 1/c = 1 \]
(24)
and with (16) and (18)
\[ b = \sqrt{1 - e^2} = \zeta \] or \( e = \sqrt{1 - \zeta^2} \).
(25)
So the meaning of \( \zeta \) is to be the ratio of the short axis to the long one. It must be chosen to lie between 0 and 1.

### 3. UHB-Treatment

We start from (11) and (12) and choose a formulation that avoids square roots, but uses products throughout instead. Basically there are two independent variables, \( \xi \) and \( \eta \), that should have the form (1) each, but the symmetry axis (\( \xi \)-axis) can be utilized to reduce the complexity. Choosing the time zero in such a way that the motion starts on a point on the \( \xi \)-axis we can write
\[ \xi(\tau) = \bar{\xi} + \sum_{j=1}^{N} \xi_j \cos(j \omega \tau) \]
(26)
and
\[ \eta(\tau) = \sum_{j=1}^{N} \eta_j \sin(j \omega \tau) \]
(27)
without loss of generality of the results, i.e. \( \xi \) is symmetric with respect to time reversal and \( \eta \) is antisymmetric, or \( \xi \) is of cosine type whereas \( \eta \) is of sine type.

This property has to be noted by application of the multiplication procedure as given in [1]. So in the numerical computation a slightly different modification of the original subroutine PROD was used that utilizes fully these simplifications with the
result of substantial time savings. In particular the result of the multiplication of two sine or two cosine variables is of type cosine, whereas mixed multiplication (cos * sin or sin * cos) yields a sine-type result. It turns out that all actually resulting — even intermediate — products are of cosine type, since \( d\eta/d\tau \) is cosine and \( d\xi/d\tau \) is sine type.

The equations most suitable for the UHB-treatment are

\[
\left( \left( \frac{d\xi}{d\tau} \right)^2 + \left( \frac{d\eta}{d\tau} \right)^2 + 1 \right) (\xi^2 + \eta^2) - 4 = 0 \quad (28)
\]

and

\[
\xi \cdot \frac{d\eta}{d\tau} - \eta \cdot \frac{d\xi}{d\tau} - \chi = 0 . \quad (29)
\]

With (26) and (27)

\[
\frac{d\xi}{d\tau} = -\omega \sum_{j=1}^{N} j \xi_{ij} \sin(j \omega \tau) \quad (30)
\]

and

\[
\frac{d\eta}{d\tau} = \omega \sum_{j=1}^{N} j \eta_{ij} \cos(j \omega \tau) . \quad (31)
\]

Forming products (particular vector products of the UHB method, special symbol \( \odot \))

\[
d = \xi \odot \frac{d\eta}{d\tau} . \quad (32)
\]

and

\[
f = \eta \odot \frac{d\xi}{d\tau} . \quad (33)
\]

we get from (29) \( N + 1 \) equations

\[
d_{ij} - f_{ij} - \chi = 0 , \quad (34)
\]

\[
d_{ij} - f_{ij} = 0 \quad \text{for} \quad j = 1 \ldots N . \quad (35)
\]

With further products

\[
g = \frac{d\xi}{d\tau} \odot \frac{d\xi}{d\tau} , \quad (36)
\]

\[
h = \frac{d\eta}{d\tau} \odot \frac{d\eta}{d\tau} , \quad (37)
\]

\[
o = \xi \odot \xi , \quad (38)
\]

\[
p = \eta \odot \eta , \quad (39)
\]

the linear functions

\[
s = g + h + 1 , \quad (40)
\]

\[
u = o + p . \quad (41)
\]

and finally the products

\[
v = s \odot s , \quad (42)
\]

\[
w = v \odot u \quad (43)
\]

we get \( N + 1 \) more equations

\[
v - 4 = 0 , \quad (44)
\]

\[
\omega_{ij} = 0 \quad \text{for} \quad j = 1, \ldots, N , \quad (45)
\]

so that a total of \( 2N + 2 \) non-linear algebraic equations is balanced by \( 2N + 2 \) unknowns, namely \( \omega, \xi, \{\xi_{ij}\}, \{\eta_{ij}\} \). The whole procedure needs the formation of \( 8 \) products \( \odot \), all results being of cosine type, which indicates that the problem is highly non-linear and by no means a "harmless" particular sample. Only for \( \chi = 1 \) the result is purely linear without overtones.

4. Comparison of UHB Results with the Analytical Solution

For a given parameter \( \chi \) the analytical solution is (vd. (13) through (16) and (23) through (25))

\[
\xi = \sqrt{1 - \chi^2} \cos \phi , \quad \eta = \chi \sin \phi , \quad \tau = \phi + \sqrt{1 - \chi^2} \sin \phi \quad (46)
\]

for \( 0 \leq \phi \leq 2\pi \), but the results of the UHB treatment are \( \xi, \eta, \{\xi_{ij}\}, \{\eta_{ij}\} \) for \( j = 1, \ldots, N \). There are two possibilities to compare both results: either \( \xi \) and \( \eta \) are computed for particular times \( \tau \), or the analytical results are expanded in a Fourier series; \( \omega \) can be compared directly. The latter method seems to be more convenient, so the Fourier coefficients of the analytical functions have to be derived.

From (26) and (27) follows

\[
\bar{\xi} = \frac{1}{T} \int_{0}^{T} \xi(\tau) d\tau , \quad (46)
\]

\[
\bar{\xi}_{ij} = \frac{2}{T} \int_{0}^{T} \xi(\tau) \cos(j \omega \tau) d\tau , \quad (47)
\]

\[
\bar{\eta}_{ij} = \frac{2}{T} \int_{0}^{T} \eta(\tau) \sin(j \omega \tau) d\tau . \quad (48)
\]

Since \( \bar{\xi}, \bar{\eta} \) and \( \tau \) are given as functions of \( \phi \), these integrals have to be transformed

\[
\bar{\xi} = \frac{1}{2\pi} \int_{0}^{2\pi} \left( \sqrt{1 - \chi^2} + \cos \phi \right) \left( 1 + \sqrt{1 - \chi^2} \cos \phi \right) d\phi , \quad (49)
\]

\[
\bar{\eta} = \frac{1}{2\pi} \int_{0}^{2\pi} \left( \sqrt{1 - \chi^2} + \cos \phi \right)^2 d\phi . \quad (49)
<table>
<thead>
<tr>
<th>( \omega )</th>
<th>UHB(_6)</th>
<th>abs. error [10(^{-2})]</th>
<th>UHB(_{12})</th>
<th>abs. error [10(^{-4})]</th>
<th>UHB(_{24})</th>
<th>abs. error [10(^{-6})]</th>
<th>exact solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.723693575315</td>
<td>0.99295523835</td>
<td>-27.63</td>
<td>0.99991953124</td>
<td>-8.05</td>
<td>1.00000000000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \xi )</td>
<td>0.613746200108</td>
<td>-4.01</td>
<td>0.65386595269</td>
<td>+1.75</td>
<td>0.65383484153</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \xi_{ij} )</td>
<td>j = 1</td>
<td>0.975695652763</td>
<td>+4.60</td>
<td>0.930481617873</td>
<td>+7.97</td>
<td>0.92968530157</td>
<td>+0.92</td>
</tr>
<tr>
<td>2</td>
<td>-0.193151567350</td>
<td>-0.18</td>
<td>-0.19239762991</td>
<td>-10.92</td>
<td>-0.19130687920</td>
<td>-1.25</td>
<td>-0.19130671725</td>
</tr>
<tr>
<td>3</td>
<td>0.048909015382</td>
<td>-1.12</td>
<td>0.05097357928</td>
<td>-2.22</td>
<td>0.05929561497</td>
<td>-0.24</td>
<td>0.05929510434</td>
</tr>
<tr>
<td>4</td>
<td>-0.013462724504</td>
<td>+0.84</td>
<td>-0.02210340531</td>
<td>-2.83</td>
<td>-0.02182057695</td>
<td>-0.33</td>
<td>-0.02182055136</td>
</tr>
<tr>
<td>5</td>
<td>0.006548897246</td>
<td>-0.23</td>
<td>0.00887186641</td>
<td>-1.11</td>
<td>0.00882952439</td>
<td>-0.12</td>
<td>0.00882946094</td>
</tr>
<tr>
<td>6</td>
<td>-0.002327952130</td>
<td>+0.15</td>
<td>-0.00389153342</td>
<td>-0.96</td>
<td>-0.00379362571</td>
<td>-0.11</td>
<td>-0.00379351920</td>
</tr>
<tr>
<td>7</td>
<td>-0.000582017052</td>
<td>-0.34</td>
<td>-0.00087617485</td>
<td>-0.040</td>
<td>-0.00078613467</td>
<td>-0.0074</td>
<td>-0.00078613464</td>
</tr>
<tr>
<td>8</td>
<td>-0.0000044773189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00002434476</td>
<td>-0.0059</td>
<td>-0.00002434468</td>
</tr>
<tr>
<td>9</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>10</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
<tr>
<td>11</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>12</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
<tr>
<td>13</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>14</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
<tr>
<td>15</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>16</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
<tr>
<td>17</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>18</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
<tr>
<td>19</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>20</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
<tr>
<td>21</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>22</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
<tr>
<td>23</td>
<td>-0.000006311874</td>
<td>-0.23</td>
<td>-0.0000192884923</td>
<td>-0.14</td>
<td>-0.0000192791696</td>
<td>-0.015</td>
<td>-0.0000192795672</td>
</tr>
<tr>
<td>24</td>
<td>-0.0000004473189</td>
<td>-0.43</td>
<td>-0.00008643332</td>
<td>-0.057</td>
<td>-0.00000434476</td>
<td>-0.0059</td>
<td>-0.00000434468</td>
</tr>
</tbody>
</table>

Table 1. Values and absolute errors of the quantities \( \omega \), \( \xi \), \( \xi_{ij} \), \( \eta_{ij} \) according to UHB\(_N\) (\( N = 6, 12, 24 \)) as compared with the exact solution.
\[ \xi_{cl} = \frac{2}{\pi} \int_0^{2\pi} \left( \sqrt{1 - \xi^2} + \cos \phi \right) \cdot \cos \left[ \left( \phi + j \sqrt{1 - \xi^2} \sin \phi \right) \right] \cdot (1 + \sqrt{1 - \xi^2} \cos \phi) \, d\phi, \quad \tag{50} \]

\[ \eta_{sl} = \frac{2}{\pi} \int_0^{2\pi} \xi \sin \phi \left( \phi + j \sqrt{1 - \xi^2} \sin \phi \right) \cdot \left( (1 + \sqrt{1 - \xi^2} \cos \phi) \right) \, d\phi. \quad \tag{51} \]

(49) can be readily solved and yields
\[ \xi = 1.5 \sqrt{1 - \xi^2}, \quad \tag{52} \]

whereas (50) and (51) lead to Bessel functions of the first kind \( J_n(z) \) [7], namely
\[ \xi_{cl} = (-1)^n \left[ 3 \sqrt{1 - \xi^2} J_0(j \sqrt{1 - \xi^2}) - 2(1 - \xi^2) J_{n+1}(j \sqrt{1 - \xi^2}) + J_{n-1}(j \sqrt{1 - \xi^2}) \right] + \frac{1}{2} \sqrt{1 - \xi^2} \left[ J_{n+2}(j \sqrt{1 - \xi^2}) - J_{n-2}(j \sqrt{1 - \xi^2}) \right], \quad \tag{53} \]

\[ \eta_{sl} = (-1)^n \left[ J_{n+1}(j \sqrt{1 - \xi^2}) - J_{n-1}(j \sqrt{1 - \xi^2}) \right] - \frac{1}{2} \sqrt{1 - \xi^2} \left[ J_{n+2}(j \sqrt{1 - \xi^2}) - J_{n-2}(j \sqrt{1 - \xi^2}) \right], \quad \tag{54} \]

with
\[ J_n(z) = \frac{\sin(n) \sum_{k=0}^\infty (-1)^k \frac{z^{2k}}{2k! (n+k)!}}, \quad \tag{55} \]

The comparison was done for \( \kappa = 0.9 \) corresponding to \( \varepsilon = 0.43589 \) (the eccentricity of the ecliptic e.g. is only 0.0167) and the UHB treatment was executed for \( N = 6, 12, \) and 24, each higher pass using the results of the preceding lower pass as a starting point. The first two passes (\( N = 6 \) and 12) were performed with the normal program using Powell’s method for the solution of the system of non-linear algebraic equations
\[ f(x) = 0. \quad \tag{56} \]

Since the transition from \( N = 12 \) to \( N = 24 \) means only a small refinement, the much simpler though normally often nonconverging Newton-Raphson method was used which needs much less storage. Since the evaluation of the Jacobian matrix and its inversion is by far the most time-consuming step, this was performed only once and the inverse Jacobian \( J^{-1} \) of the first iteration was used unchanged for the subsequent iterations \( k \) according to
\[ x^{(k+1)} = x^{(k)} - J^{-1} f^{(k)}. \quad \tag{57} \]

The results are given in Table 1 in full accuracy of the used personal computer HP 85 together with the absolute errors. To give the absolute rather than the relative errors seems to be reasonable, because to get \( \xi (\tau) \) and \( \eta (\tau) \) itself all \( \xi_{cl} \) or \( \eta_{sl} \) have to be multiplied by \( \cos (j \omega \tau) \) or \( \sin (j \omega \tau) \), respectively, which change between -1 and +1, and because \( a, b, \) and \( \xi \) are = 1 or slightly smaller. The decrease of the exact amplitudes \( \xi_{cl}/\xi_{c1} \) is \( 4.08 \times 10^{-3} \) for \( N = 6, 4.67 \times 10^{-5} \) for \( N = 12, \) and \( 1.72 \times 10^{-8} \) for \( N = 24, \) and the decrease of \( |\eta_{sl}/\eta_{s2}| \) is very similar.

5. Conclusions

The error analysis shows that in the most significant cases (\( N = 12 \) and 24; the case \( N = 6 \) is only transient) the greatest error is in \( \omega \); it follows \( \eta_{s1}, \xi, \xi_{s2}, \xi_{c1}, \eta_{s2} \) as can be seen from Table 2. The order of the magnitudes of these quantities is \( \omega, \xi_{cl}, \eta_{s1}, \xi, \xi_{s2}, \eta_{s2} \). i.e. only \( \xi_{c1} \) has another position. It is interesting that when comparing these two cases even the signs of the errors are equal and the ratios are for \( N = 12 \) to \( N = 24 \) with \( 871 \pm 10 \) relatively constant. In each column the errors de-

<table>
<thead>
<tr>
<th>( \omega )</th>
<th>( \eta_{s1} )</th>
<th>( \xi )</th>
<th>( \xi_{s2} )</th>
<th>( \xi_{c1} )</th>
<th>( \eta_{s2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( -70.45 )</td>
<td>( +52.71 )</td>
<td>( +14.93 )</td>
<td>( -10.92 )</td>
<td>( +7.97 )</td>
<td>( -7.63 )</td>
</tr>
<tr>
<td>abs. error [10^{-4}]</td>
<td>abs. error [10^{-6}]</td>
<td>abs. error [10^{-5}]</td>
<td>abs. error [10^{-5}]</td>
<td>abs. error [10^{-5}]</td>
<td>abs. error [10^{-5}]</td>
</tr>
<tr>
<td>875</td>
<td>884</td>
<td>853</td>
<td>873</td>
<td>866</td>
<td>872</td>
</tr>
<tr>
<td>ratio errors</td>
<td>( \text{UHB}<em>{24}/\text{UHB}</em>{24} )</td>
<td>( \text{UHB}<em>{24}/\text{UHB}</em>{24} )</td>
<td>( \text{UHB}<em>{24}/\text{UHB}</em>{24} )</td>
<td>( \text{UHB}<em>{24}/\text{UHB}</em>{24} )</td>
<td>( \text{UHB}<em>{24}/\text{UHB}</em>{24} )</td>
</tr>
<tr>
<td>abs. error [10^{-4}]</td>
<td>abs. error [10^{-6}]</td>
<td>abs. error [10^{-5}]</td>
<td>abs. error [10^{-5}]</td>
<td>abs. error [10^{-5}]</td>
<td>abs. error [10^{-5}]</td>
</tr>
<tr>
<td>875</td>
<td>884</td>
<td>853</td>
<td>873</td>
<td>866</td>
<td>872</td>
</tr>
</tbody>
</table>
crease with $j$ so that the highest harmonics have not only the smallest amplitudes, but also the smallest absolute error, although their relative error would be much greater. The mean error ratio is of the same order of magnitude as the decrease of the ratio of last to first amplitude from $N = 12$ to $N = 24$ (namely 2715). It could be expected that the truncation at $j = N$, i.e. the assumption that all amplitudes $A_j$ for $j > N$ are equal to zero, causes the last amplitude $A_N$ to behave in an especially unpredictable way, but this is not observed: all amplitudes fade smoothly with higher $j$. Thus the former intuitive procedure to watch the convergence of the quantities and to let the decay of amplitudes guide the choice of the final $N$ is justified, at least in this rigorously solvable example.

Acknowledgement

This work was supported by the German “Fonds der Chemischen Industrie” which furnished the personal computer HP 85 used for the computations.