Theoretical Basis for Earth Tide Analysis with the New ETERNA34-ANA-V4.0 Program

Klaus Schueller

397/75 Moo 13,Baan Bunhiranyarak, T.Nokmueang, A.Mueang Surin Province 32000, Thailand,
klaus.e.schueller@t-online.de

Abstract

The theoretical basis for the new version of the ETERNA34 program for Earth tide analysis is presented. The functional model for the least squares analysis is derived including the tidal signal and additional processes. A hypothesis-free model of higher potential degree constituents in the basic tidal wave groups is introduced. For physical regression processes transfer functions of arbitrary length can be modelled leading to frequency dependent regression coefficients and phase shifts. The least squares parameter estimation is discussed especially with respect to the impacts of window functions, leading to maximum resolution and minimum leakage least squares estimators. The condition number κ , derived from the eigenvalues of the normal equation matrix , proved to be the overall quality criterion. The stochastical model, as implemented in the new version, is explained, now being fully in accordance with least squares theory. It is emphasized that spectrum estimation of the residuals should be based on its autocovariance function. Furthermore, it is shown, how the processing of the residuals is performed by a new tool, the "High Resolution Spectral Analyser". Finally, the problem of time-variant parameters is examined and proposals are given for detection and interpretation.

Keywords: tidal analysis, ETERNA, window functions, least squares, parameter estimation, error propagation, autocovariance function, residual spectrum.

Introduction

The objective of this initiative is to acknowledge and preserve the extraordinary intellectual and technical work of my colleague and friend Prof. Dr.-Ing. habil. Hans-Georg (Schorsch) Wenzel who passed away a long time ago. Therefore, the intention is to maintain and enhance a comprehensive and sustainable platform for Earth tide analysis which will meet the requirements of the user community all over the world.

Over the past years a considerable amount of tasks has piled up which has to be tackled and solved now. As a result of recent efforts the **new version ETERNA34-ANA-V4.0** is ready to be released to interested scientists totally free of charge.

The most important features of the new **version** are:

- Enhancement of the functional model
    - Hypothesis free modelling of the higher orders of the tidal force development.
    - Modelling of additional harmonics of tidal and non-tidal origin.
    - Modelling of transfer functions of physical regression processes leading to frequency dependent regression coefficients and phase shifts
    - Comprehensive uniform polynomial model with identical coefficients for each block.
    - Deployment of window function in combination with the least squares technology for improving analysis design and interpretation.
- Redesign of the stochastical model now fully based on statistical theory
    - Frequency dependent RMS m_0 of arbitrary spectral ranges over the whole Nyquist interval, derived from the spectrum of the autocovariance function of the residuals.
    - Derivation of 95% confidence intervals for the frequency dependent RMS m_0 and all estimated parameters, now in full agreement with least squares and statistical theory.
- Information enhancements
    - Introducing the "High Resolution Spectral Analyser (HRSA)" for thoroughly analysing the residuals and estimating and presenting residual amplitudes together with their signal to noise ratios.
    - Correction of the main tidal constituent parameters for ocean influence.
    - Comparing the corrected parameters with those of different Earth models.
    - Consequent parameterization for gaining the utmost flexibility for the users of Earth tide analysis.
- Computer platforms
    - Providing an executable of the new **ETERNA34-ANA-V4.0 version** on MS Windows 7 and 8.1.
    - Support of 32- and 64-bit MS-Windows computers.
- Further maintenance and enhancement
    - Fixing detected software problems.
    - Survey and realizing of common user requirements.
    - Realizing already planned enhancements like
        - Estimating the frequency transfer functions of physical channels as performed in the HYCON method (SCHUELLER, K. 1986).
        - Built-in time-variant analysis as performed in the HYCON method

In this presentation some important concepts of statistical inference are revisited in order to provide the basis for the latest modifications and enhancements of the ETERNA program. We will refer to the modified ETERNA program as "**the new version**" throughout this presentation. Generally, no derivations or proofs of formulas will be given, when they can easily be reviewed in literature. Also, program descriptions and implementation aspects will be dealt with in a different paper, the "**ETERNA34-ANA-V4.0 USER's GUIDE**" (SCHÜLLER, K. 2014).

1. The Earth tide observation process

1.1 The sampling process in time domain

We think of a tidal observation record as the realization of the ubiquitous tidal force as input to specific instruments (gravimeter, pendulums, strain meters etc.). This tidal force signal can then be thought of being the output y(t) of a system, ideally with infinite past and future. We assume this system to be linear and comprising all features of measurement, calibration, etc.. In the following the notations for the tidal vertical component and gravimeter observations are used although the derivatives and conclusions are analogously valid for the other components.
The sampling process, i.e. the analogue-digital converter itself and the confinement of y(t) to a specific observation interval T can be abstracted by the following model:

Let the Dirac δ –function be

$$\delta(x) = \begin{cases} +\infty, & x = 0 \\ 0, & x \neq 0 \end{cases}$$

with

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1 \tag{1.1}$$

From (1.1) the so-called Dirac comb is generated as pulse train of N unity values:

$$\mathrm{i}(t_m) = \sum_{m=-\infty}^{\infty} \delta(t - m\Delta) \qquad \{m = 0,1,2, \ldots . N \tag{1.2}$$

with Δ as the **sampling interval** between two consecutive values.

Let

$$w(t) = \begin{cases} 1 & -T/2 \leq t \leq T/2 \\ & 0 \ elsewhere \end{cases} \tag{1.3}$$

be a continuous rectangular function representing the time interval T of an observation record, then

$$\mathrm{w}(t_m) = \mathrm{w}(t) \cdot \mathrm{i}(t_m) \tag{1.4}$$

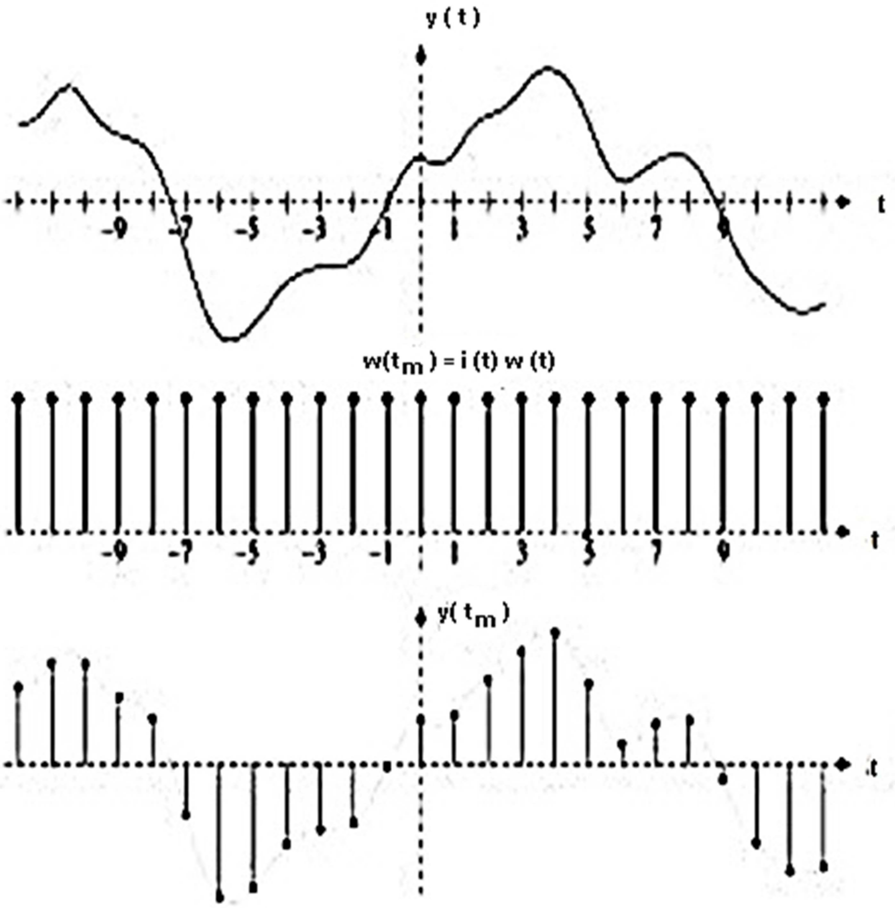
is a discrete rectangular function at discrete time points t_m in the interval T = NΔ.

The discrete observations at time t_m =mΔ within the observation period T will now be derived by (Fig.1.1) :

$$y(t_m) = \mathrm{y}(t) \cdot \mathrm{w}(t_m) \tag{1.5}$$

The function w (t_m) is known as discrete time **window function** which often is not explicitly represented in subsequent formulas. Its importance, however, will be explained in next sections.

Fig. 1: From continuous to sampled observations.



1.2 Frequency domain representation of sampled time series

Associated with the time domain, there exists a frequency domain representation for y(t), the so called "true" spectrum Y(ω). $Y(\omega)$ is continuous in frequency with an infinite frequency range. Both representations are linked by Fourier transformation:

$$y(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Y(\omega)e^{i\omega t}d\omega \tag{1.6a}$$

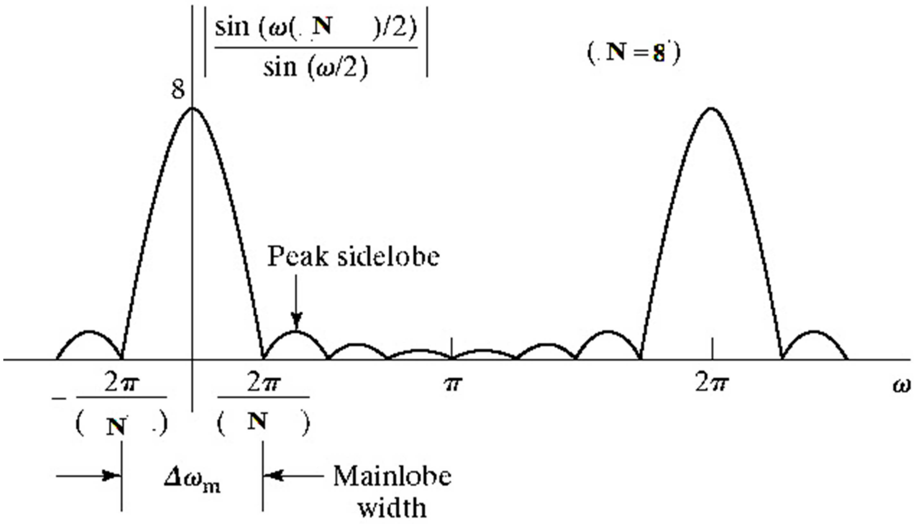
and

$$Y(\omega) = \int_{-\infty}^{\infty} y(t)e^{-i\omega t}dt \tag{1.6b}$$

Likewise, the discrete time window function $w(t_m)$ of (1.4) as sampling function possesses a frequency domain representation W(ω), the so-called spectral window function. Because $w(t_m\ )$ is a rectangular function, W(ω) can analytically be written as discrete sinc-function :

$$W_R(\omega) = \frac{\Delta}{N}\frac{\sin(\frac{\omega N\Delta}{2})}{sin(\frac{\omega\Delta}{2})} \tag{1.7}$$

Fig. 2 : sinc-function, $\Delta = 1$



Since multiplication of any two time series in the time domain means convolution of their two spectra in the frequency domain, we obtain the spectral representation $Y^{OBS}(\omega)$ of an observed time series y(t) as the **convolution** of the theoretical "true" spectrum Y(ω) with the spectral window W(ω) :

$$Y^{OBS}(\omega) = \int_{-\infty}^{\infty} y(t)w(t)e^{-i\omega t}dt$$

$$= \int_{-\infty}^{\infty} Y(\lambda)W(\omega-\lambda)d\lambda \qquad (1.8)$$

Fig. 2 and (1.8) exhibit that the spectral window $W(\omega)$ acts as a "slit" function through which we see the true spectrum Y(ω) within an uncertainty. The width of that slit is governed according to (1.7) by $T = N\Delta$, the length of the observation record. Also, we can imagine the convolution process (1.8) as bringing W(0) (the origin of $W(\omega)$) in coincidence with a specific peak of the true spectrum Y(ω) and then taking a weighted sum of Y(ω) over the whole frequency range with $W(\omega)$ as the weight function.

Conclusion:
The time window function $w(t_m)$ is fundamentally associated with the observation record because it contains all information about its frequency domain properties. The spectral window function $W(\omega)$ not only has a main lobe but is stretching over frequency with considerable side lobe peaks (see Fig. 2). Consequently, the result of the convolution at a certain frequency will be more or less a mix or smear of the "true" Y(ω) over the whole frequency domain.

1.3 General properties of window functions

It follows from (1.7) and Fig.2 that there are 2 properties which are of utmost importance when dealing with window functions, i.e.

- 1. Resolution
- 2. Side lobe convergence

Resolution is directly linked to the width of the main lobe, and is characterized by the frequency distance from the centre of the main lobe to the 1st zero position (Fig. 2). A considerable amount of window functions are offered in literature (some of them are represented in Fig. 3), but no window functions which optimally incorporate both properties.

For the rectangular window this distance is identical to the fundamental (Fourier) frequency ω_0 (1.7) which is defined as

$$\omega_0 = \frac{2\pi}{T} \quad = \frac{2\pi}{N\Delta} \tag{1.9}$$

The rectangular window stands for a window function class with optimal resolution but rather poor side lobe convergence (Fig.4,6).

The Hanning window stands for a second class of window functions with optimal side lobe convergence (Fig. 5, 6) but double frequency distance from the centre of the main lobe to the 1st zero position, i.e. $2\,\omega_0$. That means that its resolution is 2 times less the rectangular window. Its representation in the time domain is

$$w_H\ (t) = \frac{1}{2}(1 - \cos\left(\frac{2\pi}{N-1}t\right)) \qquad \text{(t= 0,.....N-1)} \tag{1.10}$$

while it's spectral window $W_H(\omega)$ can be written as a smoothed function of the rectangular window $W_R(\omega)$ (1.7) as :

$$W_H(\omega) = \frac{1}{2}W_R(\omega) + \frac{1}{4}(W_R\left(\omega - \frac{\omega_0}{2}\right) + W_R\left(\omega + \frac{\omega_0}{2}\right)) \tag{1.11}$$

Fig. 3 : Examples of window functions in the time domain

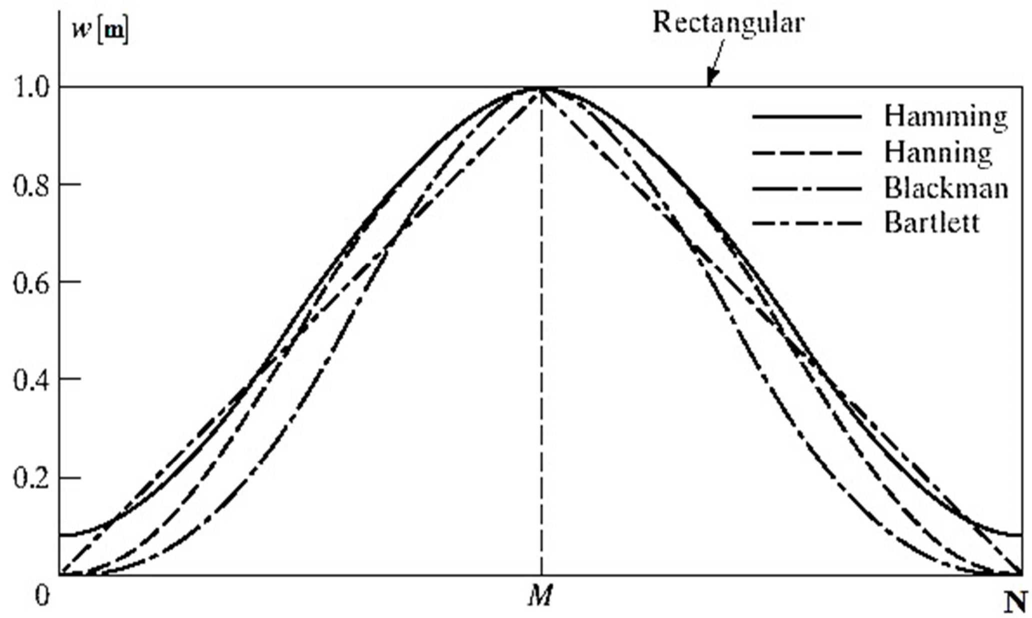


Fig. 4 : Spectral rectangular window function

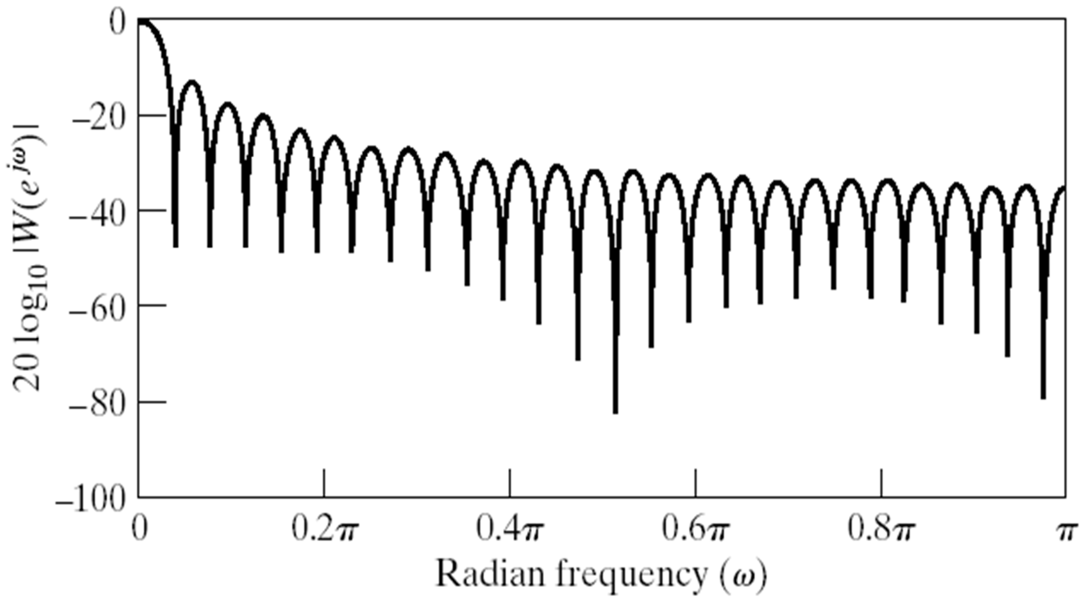
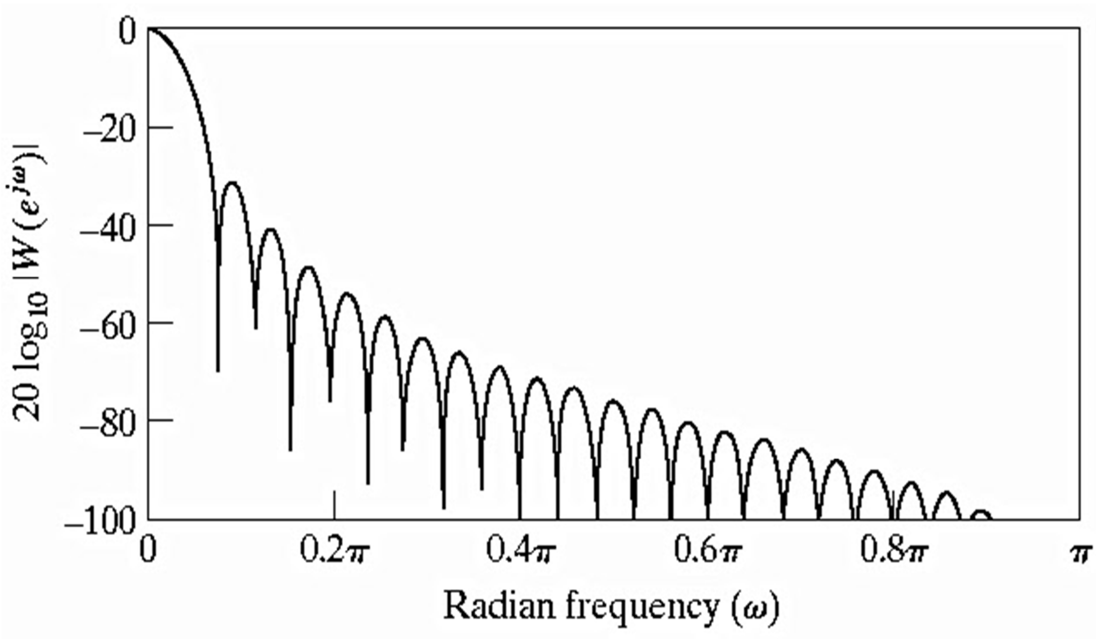


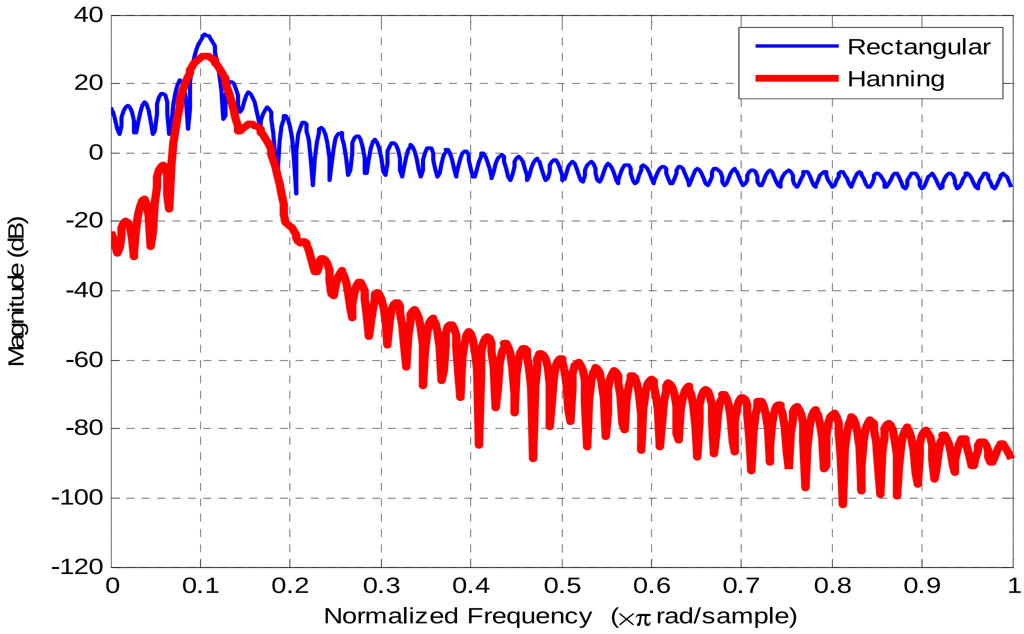
Fig. 5 : Spectral Hanning window function



From what is presented so far, it becomes quite obvious that the application of the often cited Rayleigh-criterion is nothing else than exploiting the spectral rectangular window function: for all frequencies in the observation record to be resolved, it demands a frequency distance from the centre of the main lobe to the 1st zero position of the spectral window (Fig. 2).

It will be shown later on that is criterion is by far too pessimistic, when applied in the least squares procedure.

Fig. 6: Presentation of the rectangular and Hanning spectral window function



2. The functional model for tidal observations

2.1 Tidal signal

It is well known from (Chojnicki, T. 1973), (SCHÜLLER, K. 1976), (WENZEL, H.-G. 1996) that the tidal signal $y_{ET}\ (t)$ can be modelled as

$$y_{ET}\ (t) = \sum_{i=1}^{n_{wg}} \delta_i \sum_{j=1}^{l_i} A_{ij} \cos(\omega_{ij} t + \varphi_{ij} + \kappa_i) =$$

$$\sum_{i=1}^{n_{wg}} \delta_i \cos(\kappa_i) \sum_{j=1}^{l_i} A_{ij} \cos(\omega_{ij} t + \varphi_{ij}) - \sum_{i=1}^{n_{wg}} \delta_i \sin(\kappa_i) \sum_{j=1}^{l_i} A_{ij} \sin(\omega_{ij} t + \varphi_{ij}) =$$

$$\sum_{i=1}^{n_{wg}} \delta_i \cos(\kappa_i) e_i(t) - \sum_{i=1}^{n_{wg}} \delta_i \sin(\kappa_i) f_i(t) \qquad (2.1)$$

where

- n_{wg} - number of wave groups i, i=1,..... n_{wg}
- δ_i , κ_i - tidal parameters (amplitude quotient, phase lead)
- l_i - number of tidal constituents j of wave group i, j=1,... l_i
- $A_{ij}, \omega_{ij}, \varphi_{ij}$ - theoretical amplitudes, angular velocities and phases of the i-th tidal frequency band and the j-th constituent for a rigid model Earth
- $e_i(t), f_i(t)$ - model time signals with theoretical amplitudes, angular velocities and phases of the i-th tidal frequency band introduced for abbreviation purposes

In (2.1) the underlying assumption is that the tidal parameters are constant within the n_{wg} wave groups.
However, this is principally not true, because the tidal potential is composed of different degrees l and orders m, where the orders are associated with long periodic, diurnal, semidiurnal,…etc. frequencies. The most precise development implemented in Standard ETERNA is published by (Hartmann T., Wenzel, HG. 1994) up to degree l = 6. From Earth modelling we know that the tidal amplitude factors are different for degrees 2, 3, 4, and 5 and for V2 within the different orders. Moreover, in all Earth models the phase shift κ is supposed to be 0.

The following table shows the overlapping frequency scheme of the different potential degrees:

$V_{20}\, V_{30}\, V_{40}\, V_{50}\, V_{60} \dots \dots$ – long periodic

$V_{21}\, V_{31}\, V_{41}\, V_{51}\, V_{61} \dots \dots$ – 1/1-diurnal

$V_{22}\, V_{32}\, V_{42}\, V_{52}\, V_{62} \dots \dots$ – 1/2- diurnal

$V_{33}\, V_{43}\, V_{53}\, V_{63} \dots \dots$ – 1/3-diurnal

$V_{44}\, V_{54}\, V_{64} \dots \dots$ – 1/4-diurnal

$V_{55}\, V_{65} \dots \dots$ – 1/5-diurnal

$V_{66} \dots \dots$ – 1/6-diurnal

……………………………………………………………………….

To overcome this modelling problem in (2.1), the amplitude factors of an Earth model are introduced to harmonize the heterogeneous situation within the tidal frequency bands by

$$y_{ET}\ (t) = \sum_{i=1}^{n_{wg}} \delta_i^* \sum_{j=1}^{l_i} A_{ij}^{EM}\, cos(\omega_{ij} t + \varphi_{ij} + \kappa_i) =$$

$$\sum_{i=1}^{n_{wg}} \delta_i^*\, cos(\kappa_i) \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij}\, cos(\omega_{ij} t + \varphi_{ij}) - \sum_{i=1}^{n_{wg}} \delta_i^* sin(\kappa_i) \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij}\, sin(\omega_{ij} t + \varphi_{ij})$$

$$= \sum_{i=1}^{n_{wg}} x_{c_i} \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij}\, cos(\omega_{ij} t + \varphi_{ij}) - \sum_{i=1}^{n_{wg}} x_{s_i} \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij}\, sin(\omega_{ij} t + \varphi_{ij}) \tag{2.2}$$

with

δ_{ij}^{EM} = the amplitude factors of an Earth model for each potential degree and order
A_{ij}^{EM} = Earth model amplitudes
δ_i^* = Amplitude factors between observed and model tide which is δ_i^*= 1 in case of agreement between Earth model and observations.
$\delta_i^* cos(\kappa_i)$, $\delta_i^* \sin(\kappa_i)$ -> x_{c_i}, x_{s_i} – auxiliary tidal parameters

The approach of (2.2) is equivalent to normalizing the higher potential degree amplitudes of a wave group relative to the lowest degree by the ratio higher potential degree amplitude factor to the lower one, for example: $\delta_{V3}^{EM} / \delta_{V2}^{EM}$.

(2.2) is implemented in STANDARD ETERNA, based on the Dehant-Wahr-Zschau (DWZ) model (Zschau, J. et al 1981) of a non-hydrostatic, inelastic Earth. Since wave grouping in ETERNA assumes consecutive frequencies for the different wave groups, no other modelling could be achieved, because the constituents belonging to a certain potential degree and order l,m do not exhibit consecutive frequencies but can be distributed all over the whole wave group.

In **the new version**, however, appropriate wave grouping can be automatically done by the program. The procedure comprises the following steps, each step meaning a more accurate model than the precursor step:

Step 0 : standard procedure by means of the DWZ Earth model = standard wave grouping.

Step 1: Grouping all constituents of a certain higher potential degree in a separate wave group, i.e. V3, V4, V5, V6 while the standard wave grouping refers to V20-V66.

Step2 : Grouping all constituents of a certain higher potential degree and order in a separate wave group, i.e. V30,V31,V32,V40,V41,V42,V43,V50,V51,V52,V53,V54,V60,V61,V62,V63,V64,V65, while the standard wave grouping refers to V20-V66.

Step3 : Grouping all constituents of a certain higher potential degree of a standard wave group as subgroup; for example: O1 will refer to V2, O1-3 to all V3 –constituents within the O1 group, O1-4 to all V4 –constituents within the O1-group etc..

Which step or combination of steps is adequate depends on the record length and model signal strength. It is important to emphasize that all steps can be arbitrarily combined. Hence, for each standard wave group one has to define which step should be used. This information has to be provided in the project.ini file by adding a 4-digit code for each wave group defined. This code is composed of 4 consecutive digits, one for each degree, beginning with V3, followed by V4, V5, and V6. The digits can take the values 0 to 3, meaning: 0 = step0, 1=step1, 2=step2, 3=step3.

By this procedure **the new version** is principally able to model the higher potential degrees and orders without relying on Earth model assumptions. In this context we will consider to implement the potential development of Kudryavtsev, S.M. 2004 of some 27000 constituents in **the new version.**

However, especially for weak model signals, shorter series, etc., it may be indicated to select the DWZ-Earth model information (step0) and combine it moderately with step1 (i.e. only moderate resolution requirements). Also, there will be relatively high mathematical correlations between these groups due to being fairly close together in frequency. Model calculations, however, proved that the numerical stability is guaranteed as long as the normal equation matrix is solvable.

2.2 Model enhancement by non-linear and additional harmonics

Often, the residuals of a least squares tidal analysis exhibit energy concentrations at tidal plus non-tidal frequencies. The cause of these concentrations may have various reasons, e.g.

- Oceanographic effects
    - Loading
    - Attraction of water masses
    - ………

- Meteorological influences
  - Air pressure
  - Rainfall

    ………
- Unknown influences
  - …….

The **first group** is characterised by the fact that the influencing processes are progressing with the same frequency as the Earth tides. In this case, analytical separation is not possible . Corrections can only be applied aposteri by means of load vectors from model calculations. To meet this requirement in **the new version** , loading information of different ocean models can be processed. Furthermore, in addition to the DWZ, different Earth models can be defined which will be compared to the corrected tidal parameters. Also, Melchior's amplitude ratios $\frac{\delta_{M2}}{\delta_{O1}}$ and $\frac{\delta_{O1}}{\delta_{K1}}$ are calculated and compared to the Earth models results. From all these results, conclusion can be drawn to what degree the ocean corrected parameters will fit to a specified Earth model.
In case of **non-linear loading by shallow water tides** (MERRIAM ,J.B 1995), (SCHÜLLER,K. et al 1979), loading effects in the Earth tide record can be observed at frequencies where the body tide is close to or zero at all. Since spectral estimates are comparably imprecise and usually without any error information, we want to estimate the amplitudes and phases of these non-linear **(NL)** harmonics, and also derive statistical estimates about their reliability and significance. Therefore, we calculate the frequencies of the non-linear tides according to the assumed non-linear model (usually quadratic), and introduce them together with the tidal signal as harmonics into a least squares adjustment.

The **second group** can be dealt with by monitoring these processes at the observation station with the same sampling rate as the tidal signal, and introducing this information as a regression process (section 2.4). However, if such monitoring is not available, their influence in the tidal observations cannot be identified. It may produce peaks in the residual spectrum and will be treated like the third group.

The **third group** comprises significant signals of unknown origin. The frequencies ω_m of these additional constituents **(ADCONS)** are taken from the residual spectrum or any other source of information and are fed back into the least squares adjustment.

The functional model for this enhancement can then be derived by generalizing (2.2) to

$$y_{ET}\ (t) = \sum_{i=1}^{n_{wg}} \delta_i \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij}\, cos\left(\omega_{ij} t + \varphi_{ij} + \kappa_{ij}\right) + \sum_{m=1}^{n_{add}} A_m\, cos(\omega_m t + \varphi_m). \quad (2.3)$$

where ω_m is given and $A_m and\ \varphi_m$ are unknown, and n_{add} the number of additional constituents. Similar to the tidal potential development, the n_{add} constituents are to be initially defined in a **new definition file** named **NL+ADCONST.dat**. This file serves as a memory for these constituents, and is placed in the COMMDAT directory like the tidal potential definition files. To perform a specific analysis, a selected subset of the constituents of **NL+ADCONST.dat** has to be specified in the "project".ini file, similar to the body tide wave groups.

In a subsequent least squares analysis, amplitudes and phases with respect to the 1st observation time point of the record are obtained together with their RMS-errors and confidence intervals for further treatment and interpretation. This process may be performed in several iteration cycles.

2.3 Enhancement of the drift model

2.3.1 Filters

The advantage of trend removal by filters is that their properties can mathematically be evaluated by their transfer functions which are the Fourier transforms of the filter weights.
STANDARD ETERNA provides several filters in the COMMDAT directory, which can be initialized for the analysis in project.ini definition file.
Low pass filters like Pertsev 51 or HYCON-MC-49 are provided by STANDARD ETERNA. Although Pertsev's filter is working fairly well at the low frequencies, there are significant deviations from gain values = 1 at higher frequencies. Therefore, an alternative filter also with 51 coefficients is provided in **the new version** , based on the Hanning window. This filter converges at the higher frequencies to the ideal high pass filter shape (gain values = 1). It also lets pass a significant percentage of the long periodic tides so that an analysis is possible despite of filtering the drift.(Note that the published filter *5.nlf from HYCON is in error due to a typing mistake: the 1st coefficient must carry a minus sign.)
When dealing with minute observation data, care has to be taken due to aliasing. To avoid aliasing, a band pass filter has been designed based on the Blackman-Tuckey-Window to allow proceeding with hourly values after filtering. This filter can be found in the COMMDAT directory of **the new version** as BMLPA60M.nlf.
Since filtering is done by time domain convolution of the observations with the filter weights, this operation means multiplication of their spectra in the frequency domain. Beyond the stop band of the filter, where parameter estimation is occurring there can be significant deviations from the ideal properties, i.e. gain values = 1 as described for the Pertsev filter. Consequently, the effect of filtering has to be corrected for by applying the filter gain to the estimated amplitudes provided the filter gain is not too close to 0.
STANDARD ETERNA correctly applies the gain in parameter and error estimation. However, all spectral amplitudes and related quantities of the residuals are not accounted for filtering. **In the new version** a gain correction due to filtering will be applied to all residual amplitudes provided the gain is above a predefined level ,i.e. not too close to zero. Note that such corrections can reverse filtering only to a certain degree, because they are applied to the convolved spectrum $Y^{OBS}(\omega)$ (1.8), while the filter acts on the original $Y(\omega)$.
STANDARD ETERNA does not allow for modelling long periodic (LP) -tides when using filters. However, there are situations when this is desirable. Therefore, in the **new version** we do allow modelling of LP tides although a numerical high pass filter has been applied. This can be useful, when the filter eliminates only the very low frequencies and letting the LP-tides pass with sufficient signal strength.
For meteorological channels, consistency with respect to the filter is gained when these channels are processes with the same filter before being introduced into an analysis.

2.3.2 Chebychev polynomials

Chebychev polynomials are defined as

$$\cos(n\varphi) = T_n(\cos(n\varphi)) = T_n(x), \qquad x = \cos(\varphi),\ x \in [-1,1],\ n \in N$$

with

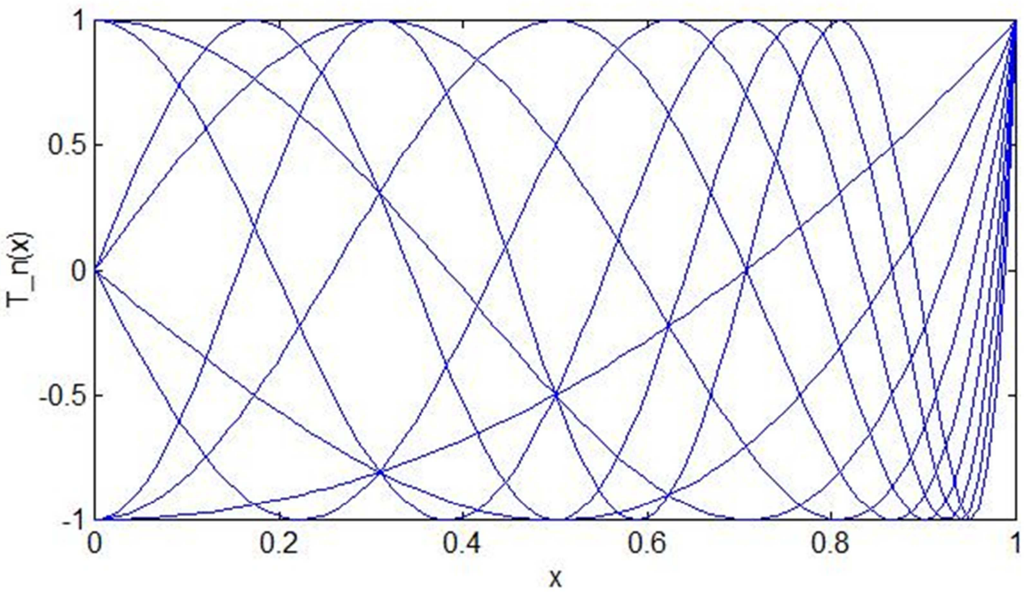
$$T_0\ (x) = 1$$
$$T_1\ (x) = x$$
$$T_2\ (x) = 2x^2 - 1$$
$$T_3\ (x) = 4x^3 - 3x$$
$$T_4\ (x) = 8x^4 - 8x^2 + 1$$
……………………………………….

where

$$|T_n\ (x)| \leq 1\ for\ x \in\ [-1\ ,1]\quad and\ n\ \in\ N \tag{2.4}$$

The first 10 polynomials are shown in Fig. 7:

Fig 7: Chebychev polynomials T_0 - T_9



The following facts about Chebychev polynomials have to be emphasized:

- the observation interval is normalized to ± 1
- the Chebychev polynomials T_i are orthogonal to each other as long as there are no gaps in the record
- other as in the case of filters, where the filter gain represents the degree of effectiveness, there is no such quantity for the Chebychev polynomials.
- a reasonable measure of the effectiveness is the minimum RMS m_0 criterion for approximations of different order combined with a statistical t- and F-Test, testing the

effectiveness of additional model parameters with a given probability.

A heuristic rule for the order u of the Chebychev polynomials to be modelled can be adopted as

$$u < \frac{N\Delta}{P_T}$$

where $N\Delta$ is the observation interval and P_T is the longest tidal period in absolute time. Otherwise, due the oscillating nature of the Chebychev polynomials (Fig.7), extremely high mathematical correlations of the polynomials with tidal (long periodic) constituents or regression processes could occur.

1.3.2.1 Block wise polynomial modelling

In case of observations with gaps STANDARD ETERNA models the polynomials for each block by a different set of polynomials coefficients. This procedure has the following disadvantages:

- the Chebychev polynomials are set up for every single block without joining conditions for the parameters over the block boundaries, so they might introduce artificial, unwanted steps into the residuals.
- each block is generally different in length so there will be different number of coefficients for each block.
- the number of polynomial coefficients will be inflated when there are many gaps in the record.

1.3.2.2 Uniform polynomial model

Since we have to assume that the observation record is homogeneous in such a sense that discontinuities are eliminated by data pre-processing**, the new version** optionally offers the definition of a polynomial over the block boundaries with one uniform parameter set. Hence, in case of applying polynomials as drift model ,the functional model (2.3) is extended as

$$y(t) = \sum_{i=1}^{n_{wg}} \delta_i \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij} \cos(\omega_{ij} t + \varphi_{ij} + \kappa_{ij}) + \sum_{m=1}^{n_{add}} A_m \cos(\omega_m t + \varphi_m) + \sum_{k=0}^{n_p} a_k T_k(x) \quad (2.5)$$

where

a_k - k =1,… n_p - polynomial coefficients also denoted as bias parameters

$T_k(x)$ - Chebychev polynomials

Moreover, in **the new version** ,the table of the bias parameters is extended by the student value

$$t_i = \frac{a_i}{m_{a_i}} \quad (2.6)$$

which is an analogue to the signal-to-noise-ratio for the spectrum in order to test the significance of the polynomial coefficients (see section 5.4).

2.4 Meteorological and other regression processes

Meteorological input channels like air pressure, ground water, temperature etc. as well as channels like the pole tide are modelled in STANDARD ETERNA by a single regression coefficient which is constant over frequency. The **new version** generalizes this model by introducing for each channel l of n_r channels a transfer function $h_l(t)$ of arbitrary length τ_l (Box, G.E.P. et al. 1994) . The total functional model of tidal observations can now be written as

$$y(t) = \sum_{i=1}^{n_{wg}} \delta_i \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij} \cos(\omega_{ij} t + \varphi_{ij} + \kappa_{ij}) + \sum_{m=1}^{n_{add}} A_m \cos(\omega_m t + \varphi_m) + \sum_{k=0}^{n_p} a_k T_k(x) + \sum_{l=1}^{n_r} \sum_{j=0}^{\tau_l} h_l(j)\, p_l(t-j) \quad (2.7)$$

where

r_l - l =1,... n_r - regression coefficients
$p_l(t)$ - physical channels (air pressure, groundwater... etc.)
$h_l(j), j = 0,..\tau_l$ - unknown transfer function weights of physical channels (air pressure, groundwater... etc.)

Fourier transforms $H_l(\omega)$ of $h_l(j)$ lead to frequency dependent regression coefficients $R_l(\omega)$ and associated phase shifts $\phi(\omega)$.

2.5 Stochastic and residual processes

Since tidal observations are obtained by measurements a stochastic component has to be taken into account which change (2.7) to

$$y(t) = \sum_{i=1}^{n_{wg}} \delta_i \sum_{j=1}^{l_i} \delta_{ij}^{EM} A_{ij} \cos(\omega_{ij} t + \varphi_{ij} + \kappa_{ij}) + \sum_{m=1}^{n_{add}} A_m \cos(\omega_m t + \varphi_m) + \sum_{k=0}^{n_p} a_k T_k(x) + \sum_{l=1}^{n_r} \sum_{j=0}^{\tau_l} h_l(j)\, p_l(t-j) + \varepsilon(t) \quad (2.8)$$

The assumption here is that the *process* $\varepsilon(t)$ is (at least an asymptotically) an ergodic and stationary process in time, i.e. the stochastic properties of $\varepsilon(t)$ are the same for the ensemble and sample space and are not dependent of absolute time. Furthermore, it is supposed to be normally distributed with mean μ and variance σ^2 as

$$E[\varepsilon(t)] = \mu = 0,$$

and

$$E[\varepsilon(t)\,\varepsilon(t)] = \sigma_{\varepsilon\varepsilon}^2$$

A process with these properties will also be referred to as “white noise”.

Even if the normal assumption is not fulfilled, we can adopt the results from the normally distributed case as approximation with respect to the Central Limit Theorem (Jenkins ,G.M. et al 1968) .
In most practical cases $\varepsilon(t)$ will not only contain stochastic parts but various kinds of residual processes , e.g. measurement errors, model errors due to inappropriately modelled signal components as well as non-modelled signals. Further examples of such contents could be :

- remaining parts of the instrumental drift,
- other additional signals due to physical reasons (e.g. ocean loading, additional unmodeled physical signals like rainfall temperature etc.),
- part time or seasonal signals like storm surge loading etc.,
- white and coloured noise .

These processes lead to $\varepsilon(t)$ not being initially random . In this case an iterative analysis procedure is suited to eliminate these contents to the highest possible degree.

2.6 Sampling intervals of tidal observations

2.6.1 Nyquist frequency

The Nyquist frequency is defined as

$$\omega_{\upsilon} = \frac{360}{2\Delta} \tag{2.9}$$

It represents the highest frequency that can uniquely be resolved for a given sampling interval Δ.
For hourly data the Nyquist frequency will be

$\omega_{\upsilon 1h} = \frac{360}{2}$ =180°/h or 12 cpd.

For minute data the Nyquist frequency will increase to

$\omega_{\upsilon 1m} = \frac{360}{2\frac{1}{60}}$ =10800°/h =720 cpd.

2.6.2 Hourly data

There is no strong argument for introducing shorter sampling intervals than 1 hour to the observation record, because STANDARD ETERNA is mainly designed for hourly data and only processes tidal frequencies in the least squares adjustment. Filtering is only possible at this sampling interval. Furthermore, the spectral analysis of the residuals is restricted to 65°/h or 4.3 cpd.
Nowadays, the trend goes to shorter sampling intervals without taking advantage of the larger Nyquist frequency interval.
In the **new version,** any frequency of the Nyquist interval can be processed in least squares as well as spectral analysis..

Despite of the fact that enormous computer power is nowadays available, there is no need for dealing with a 60 times higher amount of data, if there is no gaining of additional information from the observations.
As pointed out in section 2.3.1,if minute data are sampled and transformed to one hour sampling interval, an aliasing filtering has to be done in advance.

2.6.3 Minute data

In the new version, two filters dealing with minute samples are added and placed into the COMMDAT directory:

- interpolated Pertsev 51 filter (PERLP60M.nlf ,length =3001 min)
- filter based on the Hanning window (SCHLP60M.nlf, length = 3001min)

These filters can be applied for removing the long periodic signals from minute data.
It is planned for the new version to provide an option to perform Earth tide analysis with hourly data even if the input data set is composed of minute data. Also, for estimating the autocovariance function the minute sampling interval will be changed to an hourly one.

These two objectives, however, demand that there no significant energies at frequencies higher than 180°/h or 12 cpd in the observation record. If this assumption cannot be assured in advance, an aliasing filtering with stop band higher than 180°/h has to be processed first in order to avoid aliasing effects.
The new version provides such a filter also in the COMMDAT directory denoted by

- BMLPA60M.nlf, based on the Blackman-Tuckey window.

Its 1st part is a filter of length = 3001 min filtering out the very low frequency energies. This part is combined with a 2nd part of filter removing energies higher than 180°/h or 12 cpd by smoothing the minute observations over 361 min. All in all the filter is of length 3361 min and it is removing the long period signals with cut-off at 120°/h= 8 cpd. Applying this filter solves the problem of aliasing to a sufficient degree.

3. Parameter estimation in the least squares model

Let us define conventions first:
In statistic references there are presentations where “estimators” and “estimates” always exhibit different notations.
In this presentation when dealing with least squares analysis, “estimators” (=estimation rules) are given whenever possible in matrix notations while the “estimates” (=results of a specific analysis) will be the elements of the associated vectors or matrices. Also, the context will make clear when we are dealing with estimators or estimates.

3.1 Least squares target functions

In matrix-notation (in the following written in bold) the complete model of (2.8) can be written as (Wolf ,H. 1968):

$$\boldsymbol{v} + \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$$

or

$$\boldsymbol{v} = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{y} \tag{3.1}$$

where according to least squares convention v(t) = - $\varepsilon(t)$ and **v** being the estimator of $\boldsymbol{\varepsilon}$**.**

Applying the least squares principle to (3.1) means to minimize the (scalar) target function

$$\Omega = \boldsymbol{v}^{\boldsymbol{T}}\boldsymbol{v} = min \tag{3.2}$$

with respect to the unknown parameters **x.** (3.2) then leads then to the well-known set of normal equations :

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{N}\,\boldsymbol{x} = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}$$

or

$$\begin{bmatrix} \sum e_1(t)\, e_1(t) & \cdots & \sum e_1\,(t) f_u(t) \\ \vdots & \ddots & \vdots \\ \ldots & \cdots & \sum f_u(t)\, f_u(t) \end{bmatrix} \mathbf{x} = \begin{bmatrix} e_1(t) & \cdots & e_1(t) \\ \vdots & \ddots & \vdots \\ f_u(t) & \cdots & f_u(t) \end{bmatrix} \mathbf{y} =$$

$$\begin{bmatrix} \sum e_1(t)\, y(t) \\ \ldots \\ \sum f_u(t)\, y(t) \end{bmatrix}$$

with the estimator

$$\mathbf{x} = (\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A})^{-1}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y} \tag{3.3}$$

In (3.3) $e_i(t), g_i(t)$ are meant to comprise all model signals (2.1)-(2.7) of the function model.

The coefficients of the normal equation matrix **N** will be recognized as the auto and cross energies of the model signals, while the absolute term $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}$ of the normal equations carry the cross energies of the modelled signals and the observations.

Let us now modify (3.2) to

$$\Omega_w = \boldsymbol{v}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{v} = min \tag{3.4}$$

with the diagonal matrix **W** = **diag**$\{\boldsymbol{w}_{t_1}, \boldsymbol{w}_{t_2}, \ldots \ldots \ldots .\boldsymbol{w}_{t_N}.\}$ containing the discrete values of the window function of (1.4).

Then we will arrive at the normal equations to be

$$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{A}\,\boldsymbol{x} = \boldsymbol{N}\,\boldsymbol{x} = \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{W}\boldsymbol{y}$$

or

$$\begin{bmatrix} \sum e_1(t)\, e_1(t) w(t) & \cdots & \sum e_1\,(t) f_u(t) w(t) \\ \vdots & \ddots & \vdots \\ \ldots & \cdots & \sum f_u(t)\, f_u(t) w(t) \end{bmatrix} \mathbf{x} = \begin{bmatrix} e_1(t) w(t) & \cdots & e_1(t) w(t) \\ \vdots & \ddots & \vdots \\ f_u(t) w(t) & \cdots & f_u(t) w(t) \end{bmatrix} \mathbf{y} =$$

$$\begin{bmatrix} \sum e_1(t)\, y(t) w(t) \\ \ldots \\ \sum f_u(t)\, y(t) w(t \end{bmatrix} \tag{3.5}$$

leading to the window dependent least squares estimator for the parameters

$$\mathbf{x} = (\boldsymbol{A^T W A})^{-1} \boldsymbol{A^T W y} \tag{3.6}$$

It is obvious that when choosing w_{t_m} to be the rectangular window, then (3.6) transforms to the already derived estimator (3.3).
(3.5) shows in detail how the window function w_{t_m} is automatically introduced to the least squares parameter estimation (a detailed presentation is given in (SCHÜLLER, K. 1976)).

The residuals **v** will be obtained by inserting (3.6) into (3.1).

3.2 The impact of window functions on least squares parameter estimation

The importance of window functions in least squares tidal analysis is often underestimated because they are only implicitly involved and the impacts are not quite obvious if not presented as in (3.5).

We see 3 major items where the influence of the window functions are of real importance:

- Least squares estimation and interpretation of tidal, non-tidal and meteorological parameters
- Estimation of frequency dependent root mean square errors.
- Estimation and interpretation of the spectral properties of the residuals

Basically, windowing in a least squares adjustment comprises both the modelled and unmodeled parts of the observation record. For the modelled part resolution problem is of utmost interest. For the unmodeled part , it is leakage.
The preceding discussions in sections 1.3 and 3.1 indicate that we are dealing with the rectangular window as representative of providing maximum resolution and the Hanning window as the one of minimum leakage. Hence, it follows quite naturally that we define 2 classes of least squares estimators:

3.2.1 Maximum resolution least squares estimator

The **maximum resolution** least squares estimator will be the one using the rectangular window matrix $\boldsymbol{W_R}$

$$\begin{aligned} \boldsymbol{x_{WR}} &= (\boldsymbol{A^T W_R A})^{-1} \boldsymbol{A^T W_R y} \\ &= (\boldsymbol{A^T A})^{-1} \boldsymbol{A^T y} = \boldsymbol{N^{-1} y} \end{aligned} \tag{3.7}$$

3.2.2 Minimum leakage least squares estimator

The **minimum leakage** least squares estimator will be the one using the Hanning or similar window matrices $\boldsymbol{W_L}$ exhibiting rapidly converging side lobes:

$$\boldsymbol{x_{WL}} = (\boldsymbol{A^T W_L A})^{-1} \boldsymbol{A^T W_L y} = \boldsymbol{N_L^{-1} y} \tag{3.8}$$

In the following a set of rules will be derived to decide which one of the estimators to apply appropriately.

3.2.3 The condition number $\boldsymbol{\kappa}$ $(\boldsymbol{N})$ as criterion for frequency resolution

From (3.5) it is obvious how the window function properties are involved in the normal equations. Assume that 2 tidal frequency bands shall be resolved, which are close in frequency. The spectral window will exhibit that their frequency distance is close to the centre of the main lobe. Hence, the associated parameters become nearly linear dependent . A consequence of linear dependency of any 2 parameters in the normal equations of (3,7), (3,8) results in the normal coefficient matrix **N** being singular or close to singularity : it can either be not inverted or only with considerable loss of numerical accuracy. From linear algebra, it is known that the quality criterion which gives information to what degree a system of normal equations can be solved is the so-called **condition number $\boldsymbol{\kappa}$ $(\boldsymbol{N})$** . It is defined as the quotient of

$$\kappa\,(\boldsymbol{N}) = \frac{\lambda_{max}}{\lambda_{min}} \tag{3.9}$$

where the largest λ_{max} and smallest λ_{min} eigenvalues of **N.** Note that in case of $\lambda_{min} = 0$ the condition number $\kappa\,(\boldsymbol{N})$ becomes infinite. For this reason STANDARD ETERNA as well as and the new version compute the condition number of $\kappa\,(\boldsymbol{N})$ as criterion for numerical stability with respect to resolution.

As $\kappa\,(\boldsymbol{N})$ is only dependent on the window function, the modelled signal and not on the observations, it follows that statistical criteria (AKAIKE, F-Test etc.) may not be applied as quality criterion, because they are not involved in the resolution problem at all. Using these criteria though, means to confuse cause and effect.

3.2.4 The impact of gaps in observation records

Gaps in tidal records are important to consider because they make the interpretation of the estimated parameters more complicated due to the comparably unfavourable properties of the associated window function.

Let us assume a tidal observation record be recorded over the interval –T/2 and T/2 with centre at the overall reference epoch T_0 and containing n_l gaps. Consequently, the observation record will consist of $n_l + 1$ blocks of N_i observations. Hence, we find the associated time window function

$w_{tot}(t_m)$ (for simplicity let us assume the rectangular window with sampling interval $\Delta = 1$) as a sequence of values 1 in the block areas and 0 in the gap areas.

Then, the overall time window function $w_{tot}(t)$ covering the whole record can be thought to be composed of $n_l + 1$individual time window functions $w_{i,t0}\ (t_m)$ with individual centre reference points t_{0i} .

The associated spectral windows for each block i can then be written as

$$W_{i,t0}(\omega) \ = \frac{1}{N_i} \sum_{t=\frac{-N_i+1}{2}}^{\frac{N_i-1}{2}} w_{i,t0}\ (t)\, e^{-i\omega t} \tag{3.10}$$

Before composing the individual spectral block windows to the overall window, they have to be referenced to the central epoch T_0 :

$$W_i(\omega) = e^{i\omega(t_{oi}-T_0)} W_{i,t0}(\omega) \tag{3.11}$$

Since Fourier transformation is linear, the individual spectral block windows $W_i(\omega)$ could be summed up to form the overall spectral window and after normalization by the total number of samples, we end up with

$$W_{tot}(\omega) = \frac{1}{N} \sum_{i=1}^{l} N_i W_i(\omega) \tag{3.12}$$

or

$$W_{tot}(\omega) = \sum_{i=1}^{l} \frac{N_i}{N} e^{i\omega(t_{oi}-T_0)} W_{i,t0}(\omega) \tag{3.13}$$

Formulas (3.10- 3.13) show that

- the individual block windows $W_{i,t0}(\omega)$ exhibit different resolutions depending on the amount of samples N_i in each block i,
- the individual block windows $W_{i,t0}(\omega)$ are influencing the total window proportional to the amount of samples N_i,
- the analytical expression for the overall spectral window is a sum of sinc-functions (1.7) with in- and out-of-phase components depending on each block. Different from the case of no gaps(where we only have to deal with a single sinc- function of kind (1.7), the behaviour of this composed function is rather difficult to predict. Generally, its properties can only be evaluated numerically. Our experience shows that the resolution and side lobe properties are worse compared to a window without gaps.

Recommendation:

Gaps should be filled, whenever it is possible. The effects of inserting predicted data cannot be worse than the properties of gappy records.

When using polynomials caution has to be taken with the respect to the choice of the window function. Polynomials in a least squares procedure tend to suck up any energy they can get hold of over the whole frequency domain. The Hanning window, however, is a tool for sheltering frequency domains against leakage from more distant ones. As a consequence, the two concepts are in competition with each other. Therefore, if **polynomials are applied** in a least squares model the **rectangular window** should be preferred as window function.

3.2.5 Wave grouping

Basically, with respect to what was derived in section 2.1, wave **grouping should always occur at maximum possible resolution** (see examples in (SCHÜLLER ,K. 2014)) to avoid model insufficiencies. That means that as much wave groups as possible should be introduced to the functional model (2.1) in order to reduce the assumptions on the parameters to be estimated and so reducing model errors. Having stated this principle, it is recommended to pursue the following procedure for tidal wave grouping:

Given an observation record, one has to calculate the associated spectral window first, find the first zero pass or minimum position in frequency and then do the wave grouping according to the rules derived. Usually, a separation of half or even ¼ of that distance will lead to satisfactory results. The condition number will definitely decide whether the normal equation can cope with the resolution assumptions.

In this respect the Rayleigh-criterion, especially when dealing with least squares adjustment ,is by far too pessimistic, since it postulates a frequency difference of at least the fundamental frequency between two neighbouring harmonics to be resolved (also see Munk, W. ,Hasselmann, K. 1964). Since STANDARD ETERNA automatically eliminates tidal wave groups if the Rayleigh-criterion is not met, this elimination is dropped in **the new version**. If the resolution is supposed to be too optimistically chosen, the normal equations (3.6) can either not or only poorly be solved (see **condition number** as criterion). The user will then be notified to repeat the analysis with less resolution once again.

Modelling the higher potential degrees (see section 2.1) for m < l we face a special situation. The V_l - and V_{lm} - groups will contain harmonics of several standard tidal bands and hence the resolution between single tidal bands and multi-band groups is usually sufficiently guaranteed. However, when dealing with subgroups of a tidal band, the observation record has to be sufficiently long to keep the mutual mathematical correlations low. **The new version** provides an enhanced table of wave groups where the group amplitudes and frequencies are listed. Thus, in a first step of analysis the user can easily obtain an overview about the situation with respect to resolution. For details see (SCHÜLLER,K. 2014).

The similar principle holds for signals of non-tidal origin according to (2.3). Meteorological signals play a different role since they are multi-frequency signals. Provided , the spectrum of these signals do not contain dominant energy concentrations at the tidal frequencies, but are rather smooth over frequency, the resolution problem does not appear, since we are estimating one regression parameter ,i.e. a constant transfer function for all frequencies.

In the **new version** both the spectral rectangular and Hanning window as well as their difference is calculated and print plotted similar to Fig. 5 so that the user will obtain the required information for decision making.

For the rectangular window the required record lengths to resolve tidal bands and constituents respectively are (from the most pessimistic point of view of the Rayleigh resolution):

-	27 days ->	1 month separation distance	0.549017 °/h
-	180 days->	6 months	0.082000
-	365 days ->	12 month	0.041069
-	435 days		0.034480
-	8.8 years		0.004642
-	18.6 years		0.002206
-	20942 years		0.000002

For the Hanning window the length have to be doubled (see (1.11).

At the Geo-Observatorium Odendorf of Prof. Dr.-Ing. Manfred Bonatz, more than 10 years of observations are available which can resolve the frequency distance = 0.004642°/h equal to the Moon's perigee.

Presently, one of the longest if not the longest superconducting gravimeter record available will be of about 25) years (Calvo, M. et al. 2014) which should easily resolve a frequency distance equal to 0.002206°/h.

Efforts have been made to resolve the K1 triplet being separated by the Moon's ascending node frequency (DUCARME,B. 2011). Resolving these wave groups is nothing special because it only needs a sufficient record length to acquire the necessary resolution. As it is shown in this presentation (see (3.5)), the spectral window functions are most suited for anticipating the resolution and leakage properties of a specific observation record. Therefore, it is highly recommended for future publications to present information about the underlying window properties too. Particularly, when using analysis methods with larger sampling intervals than 1 h, its sampling properties have to be fully understood: it was shown in (SCHUELLER, K. 1978) that generally such methods are extremely sensitive to aliasing and leakage effects.

4. The stochastical model of least squares

4.1 Least squares parameter errors

The errors of the parameters are determined from (3.7),(3,8) by applying the error propagation law as:

$$\begin{aligned}\boldsymbol{m_x} &= m_0\boldsymbol{N^{-1}A^TQ_WAN^{-1}}\\ &= m_0\boldsymbol{N^{-1}}\end{aligned} \tag{4.1}$$

Note that according to (2.2) $\boldsymbol{m_x}$ for the tidal model will contain the RMS of the auxiliary unknowns x_{c_i}, x_{s_i}. To derive the RMS for the tidal parameters themselves, the error propagation law has to be applied leading to

$$m_{\delta_i^*} = \frac{m_0}{\delta_i^*}\sqrt{x_{c_i}^2 Q_{x_{c_i},x_{c_i}} + x_{s_i}^2 Q_{x_{s_i},x_{s_i}} + 2x_{c_i}x_{s_i}Q_{x_{cs_i},x_{s_i}}} \tag{4.2a}$$

where Q_{ij} are the elements of $\boldsymbol{N^{-1}}$.

With $\delta_i = \delta_i^* \, \delta_{ij}^{EM}$

it follows for any constituent j of the i-th tidal frequency band :

$$m_{\delta_i} = \delta_{ij}^{EM} m_{\delta_i^*}$$

and

$$m_{\kappa_i} = \frac{m_0}{\delta_i^{*2}} \sqrt{x_{s_i}^2 Q_{x_{c_i},x_{c_i}} + x_{c_i}^2 Q_{x_{s_i},x_{s_i}} - 2x_{c_i} x_{s_i} Q_{x_{cs_i},x_{s_i}}} \qquad (4.2b)$$

Because the Q_{ij} are determined by the functional model, special attention has to be dedicated to the appropriate estimation of the RMS m_0.

4.2 Parseval's theorem and root mean square error $\boldsymbol{m_0}$

To reveal the structure of m_0, let us consider its representation in the time and frequency domain. The relation is provided by Parseval's theorem **:**

$$\Omega = E_v = \sum_{t=0}^{N-1} v(t)v(t) = \frac{N}{2} \sum_{i=1}^{\frac{N}{2}} A^2\,(\omega_i) \qquad (4.3)$$

where $A^2(\omega_i)$ are the quadratic amplitudes of the spectrum of the residuals v(t) at integral multiples of the fundamental frequencies. The frequencies ω_i turn out to be the Fourier frequencies

$$\omega_i = \frac{2\pi}{N} i \qquad (4.4)$$

and are consequently identical with the zero positions of the associated rectangular window; therefore, the $A^2(\omega_i)$ are mutually uncorrelated.

For v(t) as **white noise**, its amplitudes $A^2 = A_{wn}^2$ are by definition constant over frequency; so we can write with u frequencies of the model signals

$$E_v = \sum_{t=1}^{N} v(t)v(t) = \frac{N}{2} \frac{(\mathrm{N-2u})}{2} A_{wn}^2 \qquad (4.5)$$

Because the parameters are already estimated at u frequencies in a least squares adjustment, there will be (N-2u)/2 frequencies left to contribute to the overall energy.

Dividing both sides by (N-2u) leads to an unbiased estimate of the mean energy which is equal to the estimated variance m_0^2 of (4.2a):

$$m_0^2 = \frac{E_v}{N-2u} = \frac{1}{N-2u} \sum_{t=1}^{N} v(t)v(t) = \frac{N}{4} A_{wn}^2 \qquad (4.6a)$$

and

$$A_{wn}^2 = 4\,\frac{m_0^2}{N}$$

or $$A_{wn} = 2\,\frac{m_0}{\sqrt{N}} \tag{4.6b}$$

The expression (4.3) is the fundamental formula for generalizing the least squares error propagation to non-white noise processes. In this case, the spectrum of the residuals is generally not constant over frequency. However, if we assume that it is fairly constant or at least smooth within certain frequency domains $d_i = \omega_{i_e} - \omega_{i_a}$ covering the whole Nyquist interval, we can rewrite (4.3) with (4.6a) to

$$m_0^2 = \frac{1}{N-u}\sum_{t=1}^{N} v(t)v(t) = \sum_{i=1}^{n_d} m_{0,i}^2 = \frac{N}{4}\sum_{i=1}^{n_d}\frac{1}{n_i-2u_i}\sum_{j=1}^{\frac{N_i}{2}} A^2(\omega_j) =$$

$$\frac{N}{4}\sum_{i=1}^{n_d}\frac{1}{N_i-2u_i}\sum_{j=1}^{\frac{N_i}{2}} A^2(\omega_j) \tag{4.7}$$

with

- n_d = number of frequency domains d_i , i =1… n_d , usually the long periodic, diurnal, semi-, ter-, quad, 12- diurnal bands, where the frequency dependent variances are estimated
- $m_{0,i}^2$ = frequency dependent variances of domain d_i ,
- $\frac{N_i}{2}$ = number of multiples of the fundamental frequency in the ith domain
- u_i = number of tidal groups in domain d_i introduced to least squares analysis
- A(ω_j) = amplitudes of the spectrum at multiples of the fundamental frequencies in domain i
- ω_{i_a} = lowest Fourier frequency of the i-th frequency domain
- ω_{i_e} = highest Fourier frequency of the i-th frequency domain
- v_i = (N_i - $2u_i$) = degrees of freedom associated with $m_{0,i}^2$

The individual variances $m_{0,i}^2$ associated with each frequency domain will then be

$$m_{0,i}^2 = \frac{N}{4}\,\frac{1}{N_i-2u_i}\sum_{j=1}^{\frac{N_i}{2}} A^2(\omega_j) = \frac{N}{4}\,A_{M_i}^2 \tag{4.8}$$

where A_{M_i} is the RMS-amplitude of the i-th domain (compare with (4.6b).

A good choice for these domains is a width of $\pm\,3.75°/h$ around 15, 30, 45, 60, 75, 90,..,176.25, 180°/h or 1 to 12 cpd respectively.

Energy averages are then taken other as in the case of white noise not over the whole Nyquist interval, but only within the bounds of these domains.

The RMS errors for the auxiliary tidal parameters $\delta_i^*\cos(\kappa_k)$ and $\delta_i^*\sin(\kappa_k)$ in the i-th domain associated are calculated according to (4.1) as

$$m_{x_k} = m_{0,i}\sqrt{Q_{x_k x_k}} \tag{4.9}$$

(4.9) holds in principle for all parameters of the functional model which can be related to one of the defined frequency domains. However, there are cases where the parameters cannot be related to a specific domain (e.g. regression signals). In this case a good approximation is to choose that $m_{0,i}$ which represents the frequency domain of the most significant model signal energies.

5. Frequency dependent least squares error propagation

The previous section emphasized the need for estimating the residual spectral amplitudes for the least squares error propagation. However, we have not yet shown by what method this could be achieved.

In STANDARD ETERNA the Fourier spectrum of the residuals is calculated for determining the residual amplitudes based upon the Fourier decomposition

$$v(t) = \frac{a_0}{2} + \sum_{i=1}^{\frac{N}{2}} \left(a_i \cos\frac{2\pi i}{N} + b_i \sin\frac{2\pi i}{N} \right)$$

with the coefficients

$$a_i = \frac{2}{N} \sum_{i=0}^{N-1} v(t_i) \cos\frac{2\pi i}{N}$$

$$b_i = \frac{2}{N} \sum_{i=1}^{N-1} v(t_i) \sin\frac{2\pi i}{N} \tag{5.1}$$

and the Fourier amplitude spectrum

$$A_i = A(\omega_i) = \sqrt{a_i^2 + b_i^2} \tag{5.2}$$

at the Fourier frequencies$\omega_i = \frac{2\pi}{N} i$.

Since v (t) is assumed to be normally distributed with $E[v(t)] = 0$ (section 2.5), it follows that the expectation values of the Fourier series coefficients a_i and b_i , i.e.

$$E[a_i] = 0, \quad E[b_i] = 0 \qquad \text{and with} \qquad A_i = A(\omega_i) \to 0 \tag{5.3}$$

the residual amplitudes are tending to zero and so do their squares. This result was also confirmed by means of generated stochastical test series and numerical experiments. Therefore, this approach will not lead to appropriate solutions in case of random time series. Instead, the approach of spectral estimation via the autocovariance function of the residuals will lead to the proper solution as it will be shown in the next sections.

5.1 The autocovariance function

5.1.1 Definition and properties of the autocovariance function

The autocovariance function $c_{\varepsilon\varepsilon}^{*}(\tau)$of the stochastic process $\varepsilon(t)$ is defined as

$$c^*_{\varepsilon\varepsilon}(\tau) = E[(v(t)-\mu)(v(t+\tau)-\mu)] \quad (5.4)$$

where $\mu = E[(\varepsilon(t)]$ is the mean and E $[*]$ the expectation operator.

(5.4) can also be written as

$$c^*_{\varepsilon\varepsilon}(\tau) = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{+T/2} \varepsilon(t)\varepsilon(t+\tau)dt \quad (5.4a)$$

From (5.4a) an unbiased estimator of the autocovariance function of a discrete time series (in our case the residuals v(t) from a least squares tidal adjustment with u^* unknown parameters) of finite length T= NΔ , (Δ= 1 for simplicity) is derived as

$$c_{vv}(\tau) = \frac{1}{N-|\tau|-u^*}\sum_{t=1}^{N-|\tau|} v(t)v(t+\tau) \quad (5.4.b)$$

As v(t) is the estimate of the "true" of the stochastic process ε(t) of (2.8), then with (4.6a) it follows that c_{vv} (0) = m_0^2 is the estimated variance from the residuals v(t) of the (Earth tides) observations y(t).

The properties of the autocovariance function can be summarized as follows

- $c_{vv}(\tau)$ also indicates how much any two observations of distance τ are correlated.
- $c_{vv}(\tau)$ is only dependent on relative time positions τ (section 2.5)
- Gaps in the observation records are no essential problem for its estimation, since there are a lot of products $v(t)v(t+\tau)$ to average so that the autocovariance function can be estimated without being considerably biased by the gappy parts of the observations.
- $c_{vv}(\tau)$ is symmetrical, i.e. $c_{vv}(\tau)$ = $c_{vv}(-\tau)$.

Further properties of the estimated autocovariance function are as follow:

- For white noise
  - $c_{vv}(\tau) = \begin{cases} c_{vv}(0) = m_0^2\ ,\ \tau = 0 \\ 0,\ \tau > 0 \end{cases}$
- Coloured noise is indicated by an autocovariance function converging to 0 as τ is increasing to a finite value τ_{conv}
  - $c_{vv}(\tau) = \begin{cases} c_{vv}(\tau) \neq 0\ \ \tau < \tau_{conv} \\ 0,\ \tau > \tau_{conv} \end{cases}$
- For periodic signals of amplitude A, frequency (ω) and phase , the auto - covariance function preserves its periodic nature for any lag τ . However, the phase information is lost, because :
  - c_{vv} (τ)= $\sum_{i=1}^{M}\frac{A_i^2}{2}\cos(\omega_i\tau)$

The autocovariance function c_{vv} (τ) is better suited than the residual process v(t) itself for classifying the nature of the residuals either to be random, deterministic or mixed-up of the two different kinds of processes. Moreover, hidden periodicities can easily be observed.

5.1.2 Length of the autocovariance function

The length n_c of the autocovariance function consists of an always **odd** number of samples, i.e.

$n_c = 2M + 1$ = N with M= $\frac{N-1}{2}$.

When calculating $c_{vv)}(\tau)$, its maximum lag could theoretically be close to T= N. However, as it should be chosen as

$$\tau_{max} = M \ , \ M \leq \frac{N-1}{2} .$$

in order to be consistent with the record length of the observations. Because the autocovariance function is symmetric, it is then of the same length N as the observation record in case N is odd ; otherwise it will one sample shorter in case N is even.

5.2 The spectrum based of the autocovariance function

5.2.1 Definition and properties

It follows from the famous Wiener-Khintchine-Theorem that autocovariance function c(τ) and spectrum $S\ (\omega)$ are related by Fourier transformation:

$$c_{vv}\ (\tau) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega) e^{i\omega t}\, d\omega$$

and

$$S\ (\omega) = \int_{-\infty}^{+\infty} \ c_{vv}\ (\tau)\ \ e^{-i\omega t}\, d\tau \qquad (5.5)$$

As c(τ) is symmetric, its Fourier transform will be symmetric as well , so $S\ (-\omega) = \ S\ (\omega)$. Moreover, the autocovariance function is discrete , so (5.5) will become

$$\mathrm{S}(\omega_i) = \ \sum_{-M}^{+M} c_{vv}\ (\tau) e^{-i\omega_i\tau} = \sum_{-M}^{+M} c_{vv}\ (\tau) \cos(\,\omega_i\tau) \qquad (5.6)$$

or with A(ω_i) being the amplitude at frequency ω_i and N=2M+1

$$\mathrm{S}(\omega_i) \ = \ \frac{N}{4} A^2(\omega_i) \qquad (5.7)$$

which is equivalent to (4.6a). (5.7) is a fundamental results because it states that the spectral decomposition of the residual variance can be derived from the Fourier transform of the autocovariance function of the residuals.

In case of white noise , the autocovariance function will consist of one value ≠ 0, i.e. c_{vv} (0). Hence, the associated spectrum values will be constant with (5.7) (see also 4.6a,b)

$$\mathrm{S}(\omega_i) \ = \ \frac{N}{4} A^2(\omega_i) = \ \text{const.} \ = \ \ c_{vv}\ (0) \ = \ m_0^2 \qquad (5.8)$$

what is completely different as derived (5.3) as Fourier spectrum of the residuals v(t)).

Since we are dealing with Fourier transforms the fundamental frequency ω_0 is likewise defined as $\omega_0 = \frac{2\pi}{n_c} = \frac{2\pi}{2M+1}$ and the harmonic frequencies are $\omega_i = \omega_0 \cdot i, \quad \{\mathrm{i} = 1, 2 \ldots \mathrm{M}.$

In **the new version** all spectral estimates either for frequency dependant error propagation or detection of additional signals are estimated by the spectrum based on the autocovariance function.

5.2.2 Spectral sampling

In STANDARD ETERNA, the frequency interval of the Fourier spectrum of the residuals is fixed to 0.25°/h. To provide a higher resolution, this interval is divided by five, resulting in an increment interval of 0.05°/h, also fixed for all length of observation records. Also, the number of spectral estimates is fixed (1300) so that the spectrum will only be calculated up to 65°/h.

In the new version the spectrum itself will be calculated at multiples of the fundamental frequency $\frac{360}{N}$ °/h divided by 5. Consequently, the resolution is dependent on the record length and there is no restriction in frequency but the Nyquist one.

5.3 Confidence intervals for frequency dependent RMS m_{0_i}

5.3.1 Degrees of Freedom

In (4.5) we have shown that the total number of degrees of freedom associated with m_0^2 is N-2u, comprising the whole Nyquist interval. It is assumed that the model parameters can be linked to a certain frequency in domain d_i. If the parameters are not of harmonic origin, only half of their numbers are to be counted in u.
The total number of degrees of freedom associated with each domain d_i can be derived with (4.7):

$$d_i = \omega_{i_e} - \omega_{i_a} \tag{5.9}$$

and $\omega_{0_i} = \frac{2\pi}{N_i}$ as

$$\nu_{d_i} = 2\frac{d_i}{\omega_{0_i}} - 2u_i \tag{5.10}$$

To obtain reliable error estimates based on a sufficient number of degrees of freedom, each of the frequency domains (5.9) must be large compared the fundamental frequency $\omega_{0_i}$of the domain so that the square residual amplitudes are averaged over a large number of samples. On the other hand, this implies that the residual spectrum has to be fairly smooth within these domains. If this is not the case the functional model must be enhanced for instance by additional model signal defined in section 2.2. This procedure can be considered as a “domain whitening process”.

Applying these rules to the residuals process v(t) of an Earth tide analysis, one has to take into account the unknown tidal parameters by subtracting 2 degrees of freedom for each tidal band from the total degrees of freedom for this domain. This principle also holds for other additional harmonics defined in the adjustment.

5.3.2 Deriving confidence intervals

It can be shown (Jenkins ,G.M. et al 1968) that the quantity $\frac{(n-u)\, m_0^2}{\sigma_{\varepsilon\varepsilon}^2}$ is distributed as χ_{n-u}^2 . Likewise, the quantity $\frac{\nu S(\omega)}{\Pi(\omega)}$ is distributed as χ_ν^2 , where $\Pi(\omega)$ is the "true" spectrum and ν the number of degrees of freedom associated with $S(\omega)$.

Then, the confidence interval parameters as lower and upper limits are derived with $\chi_\nu^2\,(\alpha) = x_\nu\,(\alpha)$ as follows (Jenkins ,G.M. et al 1968):

$$\Pr\left\{x_\nu(\tfrac{\alpha}{2}) < \frac{\nu S(\omega)}{\Pi(\omega)} < x_\nu(1-\tfrac{\alpha}{2})\right\} = 1-\alpha$$

or

$$\Pr\left\{\frac{\nu}{x_\nu(1-\frac{\alpha}{2})} < \frac{\Pi(\omega)}{S(\omega)} < \frac{\nu}{x_\nu(\frac{\alpha}{2})}\right\} = 1-\propto \tag{5.11}$$

which leads to the confidence interval for the true spectrum with probability 1-α :

$$\Pr\left\{\frac{\nu}{x_\nu(1-\frac{\alpha}{2})} S(\omega) < \Pi(\omega) < \frac{\nu}{x_\nu(\frac{\alpha}{2})} S(\omega)\right\} = 1-\propto$$

or

$$\Pr\{f_l \cdot S(\omega) < \Pi(\omega) < f_u \cdot S(\omega)\} = 1-\propto \tag{5.12}$$

The quantities f are the factors to be applied to the estimated spectrum values to determine the lower and upper bounds for the true spectral values $\Pi(\omega)$ with (1-α) –probability. Taking into account the different degrees of freedom of each frequency domain d_i , it follows:

$$f_l\,(d_i) = \frac{\nu_{d_i}}{x_{\nu_{d_i}}(1-\frac{\alpha}{2})} \quad \text{and} \quad f_u(d_i) = \frac{\nu_{d_i}}{x_{\nu_{d_i}}(\frac{\alpha}{2})} \tag{5.13}$$

Since $m_{0,i}^2$ and $S(\omega_i)$ and therefore $A^2(\omega_i)$ are related by multiplication, these factors $f_l\,(d_i)$ and $f_u(d_i)$ are valid for both the time and frequency domain (4.3). Taking the square roots of f, one obtains with (4.9) for the m_{0_i} the associated confidence intervals

$$\sqrt{f_l(d_i)\cdot}\, m_{0_i} \le c_{m_{0_i}} \le \sqrt{f_u(d_i)\cdot}\, m_{0_i} \tag{5.14}$$

These confidence intervals of m_{0_i} are calculated for each domain **in the new version**. Note that in case of filtering the gain correction is applied for presentation purposes in order to deal with the actual errors and confidence intervals.

5.4 Confidence intervals for the estimated parameters

The quantity

$$t = \frac{x_i}{m_{x_i}} \tag{5.15}$$

is distributed according to the Student's t- probability distribution function. For a given probability P and ν degrees of freedom, factors t of m_{x_i} from (4.9) can be derived so that x_i lie in the confidence interval

$$\Pr\{x_i - t_{P,\nu} m_{x_i} < x_i < x_i + t_{P,\nu} m_{x_i}\} = 1-\propto \tag{5.16}$$

Applying (5.16) to the estimated parameters, t-values are calculated for a given probability (usually 95%) and the associated degrees of freedom for each tidal domain so that the confidence intervals for the parameters can be calculated.
Note that for a large number of degrees of freedom (e.g. ν=3000 and P= 95%) the values for t will be

$$t_{95\%,3000} = 1.96 \approx 2$$

For ν=3000 and P= 68,3%, we obtain

$$t_{68.3\%,3000} = 1$$

This result means that the least squares error are special confidence intervals with t=1.

The confidence intervals for the estimated parameters of a least squares analysis are also provided **in the new version**. This t-statistic can be directly used for testing the significance of modelled parameters from 0 as zero hypothesis and hence leading to a model which is confirmed on a given probability.

5.5 Comparisons to STANDARD ETERNA error propagation

In comparison to STANDARD ETERNA we found that the tidal parameter error estimates were too high by a factor of about 1.13. A short examination of this problem leads to the cause for this deviation:

Let A_{wn} be the overall RMS amplitude of the Nyquist interval, in STANDARD ETERNA referred to as white noise amplitude , we can rewrite (4.8) as

$$m_{0,i}^2 = \frac{N}{4}\sum_{i=1}^{d_i} \frac{1}{N_i - 2u_i} \sum_{j=1}^{N_i} A^2(\omega_j) = \frac{N}{4} A_{M_i}^2 = m_0^2 \frac{A_{M_i}^2}{A_{wn}^2} \tag{5.17}$$

(5.17) is the expression used by STANDARD ETERNA and is in total accordance with the least squares principle. (Unfortunately, Ducarme,B. et al. 2006 were not aware of this context, otherwise their criticisms would have been obsolete.)

The overall white noise amplitude was calculated in STANDARD ETERNA as

$$A_{wn} = \sqrt{\pi}\,\frac{m_0}{\sqrt{N}} \tag{5.18}$$

instead of

$$A_{wn} = 2\,\frac{m_0}{\sqrt{N}} \qquad \text{(see (5.8))}$$

which leads to smaller values by a factor of 0.886. Since it is used as a corrective quantity with respect to white noise parameter errors , it appears in the denominator of (5.17) and thereby enlarging the parameter errors by $\frac{1}{0.886}$ = 1.13. The appearance of π in (5.18) is not obvious as already pointed out by (Ducarme,B. et al. 2006).

From Parseval's theorem it follows that **the averaging process of the noise amplitudes has to be quadratic,** whereas STANDARD ETERNA uses the arithmetic mean**.** Moreover, the estimates in STANDARD ETERNA are not taken at integral multiples of the fundamental frequency but at fixed frequencies, so Parseval's equation is only approximately satisfied. The amplitudes are directly estimated from the Fourier spectrum of the residual process v(t) and not by the spectrum of autocovariance function. From a numerical point of view, the results for the error estimates of the tidal parameters of STANDARD ETERNA do not differ too much, if the amplitude distributions are smooth over the frequency domains. This is not surprising because the quadratic and arithmetical means of the amplitudes are then very close.

If, however, the spectrum is considerably varying within the domains, the arithmetic mean of the amplitudes is generally smaller than the RMS one. Hence, the error estimates, being 13% too high, will be compensated by simply averaging the amplitudes. This might be one reason why the numerical results of errors from STANDARD ETERNA are surprisingly close to those derived from the exact theoretical basis.

Furthermore, STANDARD ETERNA provides a specific method of error propagation for the long periodic domain. We are of opinion that there is no reason to treat frequency bands differently.

Conclusion:

Ducarme, B. et al. 2006 concluded from all these facts that STANDARD ETERNA error estimates are not based on least squares. It was criticised that (citation) *"ETERNA RMS are, unfortunately, not least squares estimates and they depend on some intuitive assumptions"*. We do not support that statement at all. Instead we would call STANDARD ETERNA error propagation a good approximation to the least squares principle. **The new version**, however, will provide error estimation procedures which are in total agreement with the least squares principle and the rules of statistics.

6. Analysis of the residual process for additional signals

6.1 The impact of least squares adjustment on the residual process

One consequence of the least squares principle (3.2), (3.4) is that

$$\boldsymbol{A^T v = 0}$$

or

$$\boldsymbol{A^T W v = 0} \tag{6.1}$$

The meaning of these two equations is that least squares estimates parameters in such a way that the residual process is orthogonal with respect to the model signals. It follows that it is free of any information about the modelled signals.

For better understanding the impact of (6.1), let us assume the Fourier decomposition (5.1) as a simple functional model for the least squares analysis. Then, the model signals will be cosine and sine functions of time, each being the representative signal of a multiple of the fundamental frequency. In this case, the meaning of (6.1) is that the residual Fourier spectrum at these frequencies is zero, i.e. the residuals do not contain any energies at the model signals frequencies.

As the model signals for the tidal and non-tidal signals are composed of cos- and sin-functions too, the effect is similar. Therefore, no additional or at least distorted information will be gained from residual spectra at the tidal frequencies with respect to harmonics.

It follows also from (6.1) that feeding back the residuals to the least squares adjustment in place of the observations, a vector $\boldsymbol{\Delta x = 0}$ would be the result.

Therefore, let us regard the $\varepsilon(t)$ of (2.8) to be composed of 2 parts:

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon_1} + \boldsymbol{\varepsilon_2} \tag{6.2}$$

or

$$\boldsymbol{v} = -\boldsymbol{\varepsilon_1} - \boldsymbol{\varepsilon_2} \tag{6.3}$$

where $\boldsymbol{\varepsilon_1}$ will influence = disturb the tidal parameters by

$$\boldsymbol{\Delta x} = -(\boldsymbol{A^T W A})^{-1}(\boldsymbol{A^T W \varepsilon_1}) \tag{6.4}$$

Consequently, $\boldsymbol{\varepsilon_2}$ does not deliver any contribution to $\boldsymbol{\Delta x}$. This means that the residuals of a least squares adjustment only contain information about $\boldsymbol{\varepsilon_2}$ which, at the tidal frequencies, is of no great interest, because it does not do any harm to the tidal parameters.

Moreover, it follows from (6.2)-(6.4) that the estimated residual process v (t) is predominately suited **to investigate signals at non- modelled** frequencies.

6.2 High resolution spectral analysis

An important aspect is the analysis of the residual spectrum, based on the autocovariance function, for signals of non-tidal origin. In order to facilitate this purpose, **the new version** provides a new analysis tool called the "**H**igh **R**esolution **S**pectral **A**nalyser (**HRSA**)".

There, the spectrum is generated at 1/5 of the fundamental frequency all over the defined domains covering the whole Nyquist interval. In the course of these calculations, the averaging of the square amplitudes for the error calculation as well as a search for spectral peaks is performed. For each domain d_i, we obtain a mean RMS amplitude which is compared to the spectral peaks within the domain by calculating the signal to noise ratio. A signal to noise ratio > 3 is proposed to be examined in more detail for an underlying cause. Also a print plot of the domain spectrum is presented, from which one can easily observe, whether it exhibits narrow peaks or is scattering in a broadband pattern around the mean. Note that in case of filtering, the gain correction is applied.

Having isolated significant frequency locations, the functional model can be improved by feeding these frequency locations into a further least squares analysis according to (2.3), and observing ,if any improvement in model adaption is gained.

One example of an analysis might illustrate the procedure :

After having analysed a 2 years superconducting record, there were significant peaks in the N2 and L2 frequency bands left. Since the analysis was performed with the DWZ-Earth model, it could be suspected that the V3 –constituents were modelled inappropriately by the Earth model. Hence we modelled these V3 –constituents as separate groups according to step 4, section2.1.

The residual spectrum of this improved V3-model did no longer exhibit any significant peaks. Moreover, the V3-amplitude quotients from the least squares tidal analysis in the N2- and L2 band were significantly different from those of the DWZ- Earth model (such a V3 model problem was indicated by (MERRIAM, J.B 1995)).

By proceeding this way all information can be exploited from the record unless the residuals mostly contain stochastic processes. To come to that conclusion, the convergence of the autocovariance function to a δ –Funktion or at least to fade out at a certain time lag has to be observed. Equivalently, this means a convergence of the residual spectrum to a constant or at least a smooth characteristic.

6.3 Detection of tidal temporal variations in the least squares residuals

An important question which had been dealt with by several scientists is the temporal variation of the tidal parameters (e.g. Calvo, M. et al. 2014).

Since time variant tidal analysis means a considerable amount of effort, the question is, whether or not it is possible to derive this information from the spectrum of the residuals. The following theoretical consideration might contribute to the problem. Let

$$y(t) = A_{ET} cos(\omega_{ET} t + \varphi_{ET}) \tag{6.5}$$

be an Earth tide constituent, the amplitude A_{ET} of which varies with time according to

$$l(t) = A_L cos(\omega_L t + \varphi_L) \tag{6.6}$$

(6.6) is then the signal which modulates A_{ET} . The modulated Earth tide signal can then be written as

$$\begin{aligned} y(t) &= (A_{ET} + l(t))cos(\omega_{ET} t + \varphi_{ET}) \\ &= (A_{ET} + A_L cos(\omega_L t + \varphi_L))cos(\omega_{ET} t + \varphi_{ET}) \\ &= A_{ET} cos(\omega_{ET} t + \varphi_{ET}) + A_L cos(\omega_L t + \varphi_L) cos(\omega_{ET} t + \varphi_{ET}) \end{aligned}$$

or

$$\begin{aligned} y(t) &= A_{ET} cos(\omega_{ET} t + \varphi_{ET}) \\ &+ \frac{A_L}{2}\cos\big((\omega_{ET} - \omega_L)t + (\varphi_{ET} - \varphi_L)\big) + \frac{A_L}{2}\cos\big((\omega_{ET} + \omega_L)t + (\varphi_{ET} + \varphi_L)\big) \end{aligned} \tag{6.7}$$

This is an important formula, since it exhibits the occurrence of (tidal) amplitude modulation as the presence of a main lobe peak and 2 symmetrical side lobe peaks of frequency distance ω_L.

An example might illustrate, how an interpretation of additional observed phenomena could be done. Assume that the amplitude of the **ocean tide** M2 is changing slightly for instance with yearly frequency and amplitude A_L . (6.7) shows that in this case, peaks of $\frac{A_L}{2}$ are folded around M2 at $\omega_{M2} \mp 0.04°/h$. With a record length of 1 year, the side lobe peaks of equal height $\frac{A_L}{2}$ should be detected in the residual spectrum . To assure the significance of these side peaks, they should be introduced into a least squares adjustment.
With shorter record length, the side lobe peaks will lead to abnormal M2 parameters and could partly be observed in the residual amplitude spectrum. As it was shown in (SCHÜLLER, K. 1976), the subsequent time-variant tidal (auxiliary) parameters of M2 (called parameter functions in SCHUELLER, K. 1986), derived from analyses with a shifted basic interval, will oscillate with $\Delta\omega = 0.04°/h$.

If (6.6) does not only contain a single harmonic, but a series with a broad spectrum, (6.7) would exhibit not only 2 symmetric peaks but a fanned out pattern of peaks around the modulated signal. This pattern would be reflected in the residual spectrum as well. All frequencies with significant peaks would then lead to a set of additional harmonics (2.3) to model this rather complicated process appropriately.

7. Conclusions and outlook

The basic principles of least squares tidal analysis and spectral analysis of the residuals have been presented as the theoretical background for the enhancements of new ETERNA34-ANA-V4.0 program. Implementation aspects and the program description have been kept to a minimum and are published in a different presentation, the user's guide (SCHÜLLER,K. 2014).

The program itself is ready for distribution to interested parties by beginning of the year 2015. It is available upon request as executable on Windows 7, 32- and 64 bit, and Windows 8.1 platforms totally free of charge.

Acknowledgements

Such a project cannot be set up without helping hands from a great variety of supporters. I am sure that I will miss somebody, and therefore I am asking for apologizes in advance.

Special thanks owed to Mrs. Marion Wenzel, who kindly supports this initiative by authorizing me to further maintain and enhance the programs of her late husband.

The FORTRAN source code of the ETERNA programs together with a comprehensive amount of sample data was provided by the ICET, Tahiti. Moreover, the ICET was a great help to get this presentation published in the BIM very rapidly. Special thanks are owed to Prof. Dr. Jean-Pierre Barriot for his kind support.

Also I would like to express my thanks to Dr. Walter Zürn, BFO Schiltach, who dedicated a lot of his precious time for providing overall encouragement and help as well as scientific consultancy and a data set for testing purposes.

My special thanks are dedicated to Prof. Bernard Ducarme for thoroughly reviewing the manuscript of this paper. His valuable remarks contributed to improve the quality of this presentation considerably.

It was a real pleasure to get in contact again with my teammate from Bonn University, Dr. Herbert Wilmes, BKG Frankfurt, who confirmed me to proceed on this project. Dr. Wziontek, BKG, was a challenging partner in scientific discussions and provided data sets for testing, thank you for that.

My sincere thanks are owed to my highly respected teacher, PhD -supervisor and friend Prof. Dr.-Ing. Manfred Bonatz, GeoOberservatorium Odendorf, who gave so many encouragements and valuable advises for getting started with this project. He also provided all of his data for testing. I am really grateful and proud that after more than a generation of pursuing different fields of activity we are working together again as a team that apparently has never been separated. Since he was the first who installed and tested the new software he chose the hard way of our cooperation.

Last but not least I would like to express my deepest thanks to my wife Arunee for her loving patience and endless support while her husband was caught in chasing some thought or code problems.

Thanks a lot to you all again.

References:

Box,G.E.P. et al. : Time series analysis. Forecasting and Control. Prince Hall Inc., New Jersey 1994. ISBN 0-13-060774-6.

Chojnicki, T. : Ein Verfahren zur Erdgezeitenanalyse in Anlehnung an das Prinzip der kleinsten Quadrate. Mitteilungen aus dem Institut fuer Theoretische Geodaesie der Universitaet Bonn Nr. 15, Bonn 1973.

Calvo, M. et al. : Time stability of spring and superconducting gravimeters through the analysis of very long gravity records. J. Geodyn .2014

Ducarme, B.,Vandercoilden,L., Venedikov,A.P. : Estimation of the precision by the tidal analysis programs ETERNA and VAV. Bulletin d'Informations Marees Terrestres, vol. 147, Bruxelles 2011.

Ducarme, B. : The K1 triplet: Can Lunar nodal waves contribute to th study of the Free Core Nutation (FCN)? Bulletin d'Informations Marees Terrestres, vol. 141, Bruxelles 2006.

Hartmann, T., Wenzel, H.-G. : Catalogue of the earth tide generating potential due to the planets. Bulletin d'Informations Marees Terrestres, vol. 119, Bruxelles 1994.

Jenkins, G.M., Watts, D.G. : Spectral analysis and its applications. Emerson-Adam Press,Inc., Boca Raton , Florida,USA.1968. ISBN 1-892803-03-8.

Kudryavtsev, S.M. : Improved harmonic development of the Earth tide-generating potential. Journal of Geodesy (2004) 77.

Merriam, J.B. : Non-linear tides observed with the superconducting gravimeter. Geophy. J .Int. (1995) 123.

Munk, W., Hasselmann, K. :Super-resolution of tides. Studies on Oceanography 1964.

Schüller, K. : Ein Beitrag zur Auswertung von Erdgezeitenregistrierungen. Deutsche Geodätische Kommission, Reihe C Heft Nr. 227, München 1976.

Schüller, K. : About the sensitivity of the Venedikov tidal parameter estimates to leakage effects. Bulletin d'Informations Marees Terrestres , vol. 81, Bruxelles 1978.

Schüller, K., Wenzel, H.-G., Zürn, W.: Analysis of simulated nonlinear systems driven by tides. Deutsche Geodätische Kommission, Reihe B Heft Nr. 231, München 1979.

Schueller, K.: Simultaneous tidal and multi-channel input analysis as implemented in the HYCON-method. Proceedings 10th International Symposium on Earth Tides, Madrid 1986.

Schüller, K.: USER’S GUIDE ETERNA34-ANA-V4.0, Installation Kit ETERNA34-ANA-V4.0, Surin 2014

Wenzel, H.-G. : The nanogal software: Earth tide data processing package ETERNA 3.30. Bulletin d'Informations Marees Terrestres, vol. 124, Bruxelles 1996.

Wolf, H. : Ausgleichungsrechnung nach der Methode der kleinsten Quadrate. Ferdinand Dümmlers Verlag Bonn. 1968.

Zschau, J., R. Wang (1981): Imperfect elasticity in the Earth's mantle. Implications for Earth tides and long period deformations. Proceedings of the 9th International Symposium on Earth Tides, New York 1981,editor J.T. Kuo, Schweizerbartsche Verlagsbuchhandlung, Stuttgart 1981.