

Some properties of the Chebyshev approximation

prepared by: F. Mignard
affiliation : OCA
reference: GAIA-CA-TN-OCA-FM-042-1
issue: 1
revision: 1
date: 2011-02-14
status: Issued

Abstract

This note deals with the general properties of polynomial approximations and interpolations with a particular emphasis on the Chebyshev representation. It is primarily driven by the Gaia needs and the exchanges within the REMAT group, although many results have a wider range of applications. Results are given on how to determine the accuracy of an approximation and set a trade off between the degree and the size of the basis time interval. The special case of Gauss-Lobatto integration is investigated in detail as a means to ensure continuity of the representation at the boundary of each sub-interval, without serious loss of accuracy in the approximation. Finally it is shown on the basis of a realistic example that, for accurate representation of ephemeris, the cubic splines are not match to the Chebyshev approximation, for comparable volume of fitted coefficients. Practical recommendations for the representations of Solar System ephemerides are given in the conclusion.

Document History

Issue	Revision	Date	Author	Comment
1	0	2011-02-14	FM	Polyn. second kind introduced - UB comment
1	0	2010-12-10	FM	Draft completed
1	0	2010-09-10	FM	Start of the document

Contents

1 Polynomial approximations **10**

1.1 Taylor expansion 10

1.2 Least squares approximation 12

1.3 Legendre expansion 13

1.4 Chebyshev expansion 14

1.4.1 Straight expansion 14

1.4.2 Relation to standard Least Squares 15

1.5 Best approximation 16

2 Chebyshev representation **18**

2.1 Basic definition 18

2.2 Chebyshev polynomials of first kind 19

2.3 Chebyshev polynomials of second kind 22

2.4 Explicit forms 25

3 Accuracy of a Chebyshev expansion **26**

3.1 General case 26

3.2 Quasi periodic functions 28

4 Computation of the Chebyshev coefficients **32**

4.1 Principles 32

4.2 Relationship to interpolation 33

4.3 Chebyshev approximation by interpolation 35

5 Error in the numerical quadrature 42

5.1 Gauss-Chebyshev quadrature 42

5.1.1 Analytical expression for the coefficients 42

5.1.2 Error in the truncated expansion 46

5.2 Composite trapezium quadrature 47

5.2.1 Error in the truncated expansion 51

5.3 A worst case function 53

6 Numerical illustrations 58

6.1 Interpolation at Chebyshev points 58

6.2 Interpolation at Lobatto points 60

6.3 Least square fitting 60

7 Comparison to splines 63

7.1 Comparison within a simple interval 63

7.2 Comparison within a composite interval 65

8 Conclusions 70

8.1 Terminology 70

8.2 Algorithm for the Lobatto interpolation 70

8.3 Future work 72

9 References **73**

Introduction

Polynomial approximation used to represent and evaluate a continuous function is a tool appearing in many places in the Gaia data processing. Many of the calibration models in astrometry or photometry are expressed in this form with the coefficients determined with a least squares fitting on dedicated observations. The attitude determined from observations at discrete times is directly solved with quaternions being represented with a cubic spline going through a set of nodes. The solar system ephemeris and that of Gaia (at least during the preparatory phase) are all computed from a Chebyshev representation of a numerical integration or analytical expression turned into a more computing-efficient form with polynomials. Several other usages of the Chebyshev representations have also been considered within the activities of the REMAT group, for example to relate the various timescales to each others. Having a DPAC note summarising the properties, good or not so good, of the difference representations should therefore be a useful contribution extending beyond the mere Gaia group.

In its form, the present note is in fact an outgrowth of an issue raised within the REMAT group about the continuity of the ephemeris representation with Chebyshev polynomials, or more generally of any Chebyshev expansions using several sub-intervals, at the crossing of each sub-interval. In the plain Chebyshev expansion, the coefficients are computed over each sub-interval, independently of the previous or the next interval, without any constraint set to ensure the continuity. From the perspective of the representation, the discontinuity is typically of the same order as the error made by approximating the function by a polynomial over the two adjacent intervals, and this can be considered as statistically acceptable, as long as the representation is used just to evaluate the values of the function. However, solar system ephemeris are used very often as input to compute the gravitational forces acting on a minor body to feed a numerical integrator. Given the fact that quite many numerical schemes compute internally first or higher differences, numerical instabilities may occur at discontinuity point. Here the numerical differentiation is not reliable if the difference $f(x_{k+1}) - f(x_k)$ takes the two points on each side of the discontinuity. It was then interesting to see how one could deviate slightly from the standard Chebyshev expansion to ensure the continuity, without loosing too much of the remarkable properties which make the Chebyshev expansion so desirable.

The second motivation to undertake this work stems for the fact that for many years I found quite unsatisfactory of not having well grounded rules to evaluate the accuracy of a representation as a function of the polynomial degree and of the length of the sub-intervals and then no theoretical way to decide on the degree required to achieve a given accuracy. This fact was very frustrating and I never found in textbooks a criterion that could have soothed my disappointment. During

this investigation I contrived some easy-to-use quantitative rules applying to a large class of functions. They cover specifically the practical needs for the ephemeris construction and allow to control the accuracy with something more general than the magnitude of the first neglected coefficient. At least these criteria bring a justification to this widely used rule.

This note is broken down in several sections with more or less independent scopes, but with a common set of notations. In the first section I introduce on the basis of a concrete example the different meanings between the related notions of interpolation, least squares fitting and polynomial approximation and their actual representation with Legendre or Chebyshev polynomials. Then in the subsequent sections I focus on the Chebyshev approximation and derive general properties of the coefficients determined by numerical quadrature. The important issue of the accuracy of a Chebyshev approximation is dealt with in Sect. 3, with a special emphasis on the representation of the quasi-periodic functions so ubiquitous in celestial mechanics.

The continuity issue is taken up in Sect. 4 from the relationship between the Chebyshev expansion and the interpolation at a well chosen set of points (the Chebyshev-Lobatto points) and I show that the precision is never degraded by more than a factor two in the worst case. This interpolation offers an alternative solution to the proposal made by S. Klioner (SK-012) based on a constrained least squares fitting, which might be required when sample points are fixed with no possible interpolation between them.

Then practical numerical methods to compute the Chebyshev coefficients are presented in Sect. 5 with an estimation of the error with respect to the theoretical coefficients. Section 6 provides several numerical illustrations of the features introduced in the more theoretical parts. The final section (Sect. 7) is devoted to a comparison between the Chebyshev expansion and the cubic spline approximation, primarily on the basis of a numerical example.

Notations

TABLE 0.2: Various notations used throughout this document

Symbol	Meaning
$\mathcal{C}[a,b]$	Set of real-valued continuous functions on $[a,b]$
$\mathcal{C}^n[a,b]$	Set of real-valued functions n times differentiable on $[a,b]$, with $f^{(n)}$ continuous
$\mathcal{P}_n$	Set of polynomials of degree n
π_n	Set of polynomials of degree n with leading coefficient equal to 1
$\mathcal{R}(z)$	Real part of the complex number z
$\|\|\mathbf{x}\|\|_\infty$	Uniform norm (Eq. (1.3))
$\|\|\mathbf{x}\|\|_2$	Quadratic norm (Eq. (1.4))
$T_n(x)$	Chebyshev polynomial of first kind of degree n
$U_n(x)$	Chebyshev polynomial of second kind of degree n
$P_n(x)$	Legendre polynomial of degree n
$\xi_k^{(n)}$	kth root of the Chebyshev polynomial $T_n(x)$; $T_n(\xi_k^{(n)}) = 0$
$\zeta_k^{(n)}$	kth extreme value of the Chebyshev polynomial $T_n(x)$; $T_n(\zeta_k^{(n)}) = \pm 1$
c_j	Theoretical coefficient of $T_j(x)$ in a Chebyshev (resp. Legendre) expansion of $f(x)$
C_j	Approximate value of the Coefficient of $T_j(x)$ is a Chebyshev (resp. Legendre) expansion
$\phi(x)$	Result of a truncated Chebyshev (resp. Legendre) expansion of a function $f(x)$
$\sum_{i=0}^{i=n} x_i$	$= x_0 + x_1 + \cdots + x_n$
$\sum_{0}^{n}{}' x_i$	$= \frac{1}{2}x_0 + x_1 + \cdots + x_n$
$\sum_{0}^{n}{}'' x_i$	$= \frac{1}{2}x_0 + x_1 + \cdots + \frac{1}{2}x_n$

1 Polynomial approximations

In this introductory section, built around a worked example, I wish to stress the difference between what must be called a *polynomial approximation* and a *polynomial representation*. The former is the use of a polynomial of a certain degree n to approximate a function using some merit criterion to select the best approximation among a the set $\mathcal{P}_n$ of polynomials of degree n. The latter is a practical way to write these polynomials on a finite set of basis functions. The basis functions can be the canonical basis $\{1, x, \cdots, x^n\}$ or any other set of $n+1$ independent polynomials, like the Legendre or Chebyshev polynomials. The difference is illustrated with a continuous function expressed in polynomial form using various approximation criteria to show how this impacts on the distribution of the final error.

Let $f(x) = \sin(\pi x)$ a real-valued function with $1 \leq x \leq 1$. It is approximated below on $\mathcal{P}_3$, using various criteria to find $p_3(x)$.

1.1 Taylor expansion

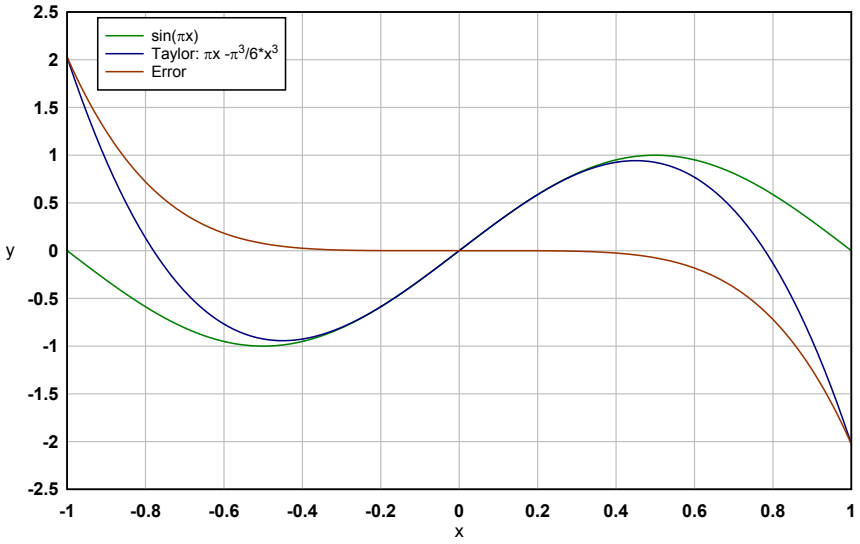


FIGURE 1.1: Taylor approximation to third degree of $\sin(\pi x)$ with $-1 \leq x \leq 1$. The plot gives the function (green), the Taylor expansion (blue) and the error (red).

To third degree on has the Taylor expansion of $f(x)$ as,

$$t_3(x) = \pi\, x - \frac{\pi^3}{6}\, x^3 \approx 3.1416\, x - 5.1617\, x^3 \tag{1.1}$$

and one has,

$$||f(x) - t_3(x)||_\infty = 2.025 \tag{1.2a}$$

$$||f(x) - t_3(x)||_2 = 0.895 \tag{1.2b}$$

where throughout this note,

$$||f||_\infty = \max_{-1 \leq x \leq 1} |f(x)| \qquad f \in \mathcal{C}[-1, 1] \tag{1.3}$$

and

$$||f||_2 = \left(\int_{-1}^{1} f^2(x)\, dx \right)^{1/2} \tag{1.4}$$

or equivalently for their discretised forms,

$$||f||_\infty = \max_{1 \leqq j \leqq n} |f(x_j)| \tag{1.5}$$

and

$$||f||_2 = \left[\sum_{j=1}^{j=n} f^2(x_j) \right]^{1/2} \tag{1.6}$$

Now, any power x^k of x can be expressed as a sum of Legendre of Chebyshev polynomials of degree $\leqq n$, meaning that (1.1) can be rewritten as,

$$t_3(x) = \pi \left(1 - \frac{\pi^2}{8} \right) T_1(x) - \frac{\pi^3}{24} T_3(x) \approx -0.7342\, T_1(x) - 1.2919\, T_3(x) \tag{1.7}$$

or

$$t_3(x) = \pi \left(1 - \frac{\pi^2}{10} \right) P_1(x) - \frac{\pi^3}{15} P_3(x) \approx 0.0409\, P_1(x) - 2.0671\, P_3(x) \tag{1.8}$$

which I call a Chebyshev or Legendre **representation** of a polynomial. Despite the appearances, this remains a Taylor expansion, with the same quality and weakness, and in no case should be mistaken for an expansion in the Fourier sense of the function $\sin(\pi\, x)$ with $-1 \leqq x \leqq 1$. As a consequence, if we are given a set of Chebyshev coefficients to evaluate a function $f(x)$, it does not necessarily mean that they have been determined with the proper criterion (discussed hereafter) that makes the Chebyshev expansion so valuable.

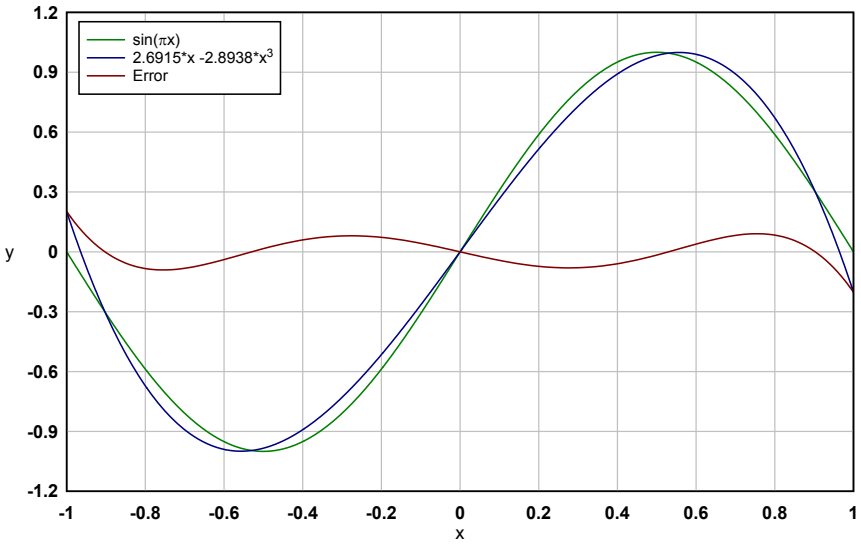


FIGURE 1.2: Least squares polynomial fitting of $\sin(\pi x)$ with $-1 \leq x \leq 1$. The plot gives the function (green), the fit (blue) and the error (red).

1.2 Least squares approximation

I consider now the polynomial fitting of $f(x) = \sin(\pi x)$ by an element $q(x) \in \mathcal{P}_3$, such that

$$||f - q||_2 \leq ||f - p||_2 \tag{1.9}$$

for any $p \in \mathcal{P}_3$. This is the usual least squares polynomial approximation.

One finds with a regular sampling of step 0.001 over $-1 \leqq x \leqq 1$,

$$q_3(x) \approx 2.6915\, x - 2.893\, x^3 \tag{1.10}$$

and one has,

$$||f(x) - q_3(x)||_\infty = 0.202 \tag{1.11a}$$

$$||f(x) - q_3(x)||_2 = 0.0937 \tag{1.11b}$$

which for both criteria is much better than the Taylor approximation of (1.2).

As before, the same polynomial can be expressed in Chebyshev or Legendre form,

$$q_3(x) \approx 0.52116\, T_1(x) - 0.72345, T_3(x) \tag{1.12}$$

or

$$q_3(x) \approx 0.9552\, P_1(x) - 1.1575\, P_3(x) \tag{1.13}$$

Remark: The polynomial fitting on a discrete representation of $f(x)$ depends on the actual sampling, so the reader doing the same application may end up with slightly different numbers.

1.3 Legendre expansion

Given the completeness of the Legendre polynomials, any continuous and square-integrable function $f(x)$ on $[-1, 1]$ can be expanded into Legendre polynomials as

$$f(x) = \sum_0^\infty c_j P_j(x) \tag{1.14}$$

where the c_j are given by,

$$c_j = \frac{2n+1}{2} \int_{-1}^{1} f(x) P_j(x) dx \tag{1.15}$$

Then one has approximately the polynomial approximation of degree k,

$$f(x) \approx \phi(x) = \sum_0^k c_j P_j(x) \tag{1.16}$$

Since the Fourier coefficient c_j in (1.15) results from a minimum of the quadratic norm, the Legendre expansion is strictly equivalent to a least squares fitting of $f(x)$ with a **regular** grid $x_i, \quad i = 1, \cdots m$, when $m \to \infty$. With the same $\sin(\pi\, x)$ as before and using an adaptative numerical procedure to evaluate the integrals in (1.15), one gets,

$$\phi_3(x) \approx 0.95493\, P_1(x) - 1.1582\, P_3(x) \tag{1.17}$$

almost identical to (1.13). The slight difference comes just from the finite sampling.

1.4 Chebyshev expansion

1.4.1 Straight expansion

This is very similar to the Legendre expansion, but for the weight function $w(x)$ in the definition of the inner product. For the Legendre system one has $w(x) = 1$, $-1 \leq x \leq 1$, while $w(x) = (1 - x^2)^{-1/2}$, $-1 \leq x \leq 1$ for the Chebyshev polynomials. Any continuous and square-integrable function can be projected on a Chebyshev base as,

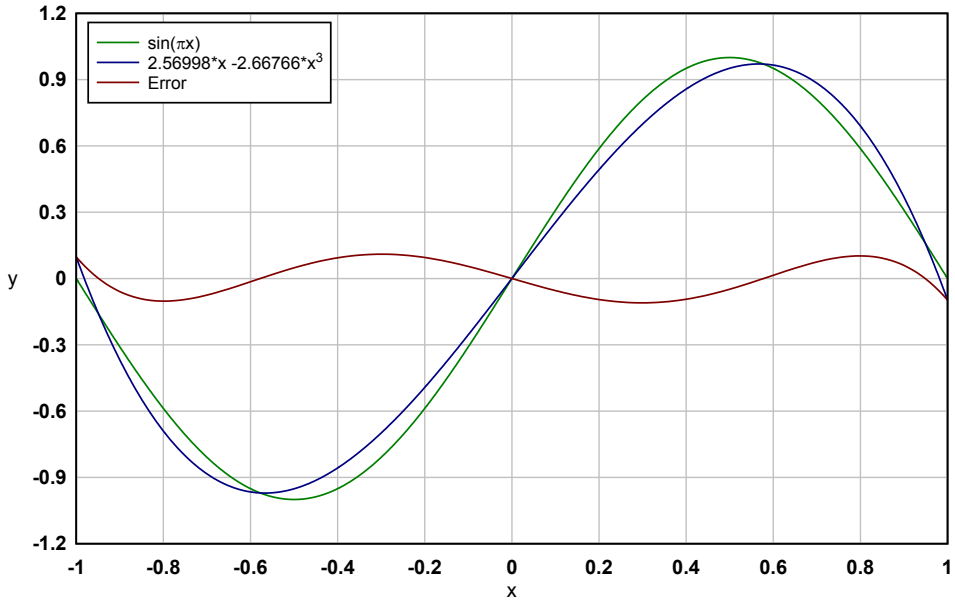


FIGURE 1.3: Chebyshev approximation, i.e truncated expansion to 3rd degree, of $\sin(\pi x)$ with $-1 \leq x \leq 1$. The plot gives the function (green), the fit (blue) and the error (red).

$$f(x) = \sum_{0}^{\infty} c_j T_j(x) \tag{1.18}$$

where the c_j are given by,

$$c_j = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x) T_j(x)}{(1 - x^2)^{1/2}} dx \tag{1.19}$$

Here one gets with a numerical quadrature,

$$\phi_3(x) \approx 0.569231\, T_1(x) - 0.666917\, T_3(x) \tag{1.20}$$

or equivalently with $T_1(x) = x$ and $T_3(x) = 4x^3 - 3x$,

$$\phi_3(x) \approx 2.569981\, x - 2.66766\, x^3 \tag{1.21}$$

and eventually,

$$||f(x) - \phi_3(x)||_\infty = 0.110 \tag{1.22a}$$

$$||f(x) - \phi_3(x)||_2 = 0.106 \tag{1.22b}$$

One sees the very large improvement in uniform norm compared to the Legendre (equiv. unweighed least squares) expansion in (1.11). The largest error in absolute value is twice as small with the Chebyshev expansion. Naturally, the error measured by the least squares norm is slightly larger since the polynomials (1.20) or (1.21) are not the best least squares fit with uniform weight.

1.4.2 Relation to standard Least Squares

The relationship between the Chebyshev expansion defined by (1.18)-(1.19) and the least squares approximation is an important issue for numerical computation. The Chebyshev polynomial is the best weighted least squares fit with the weight $w(x) = (1 - x^2)^{-1/2}$ and a discrete version can be considered with a uniform sampling $(x_1, x_2, \cdots x_n)$ over $[-1, 1)$ and a weighted least squares fit with $w(x_i) = (1 - x_i^2)^{-1/2}$. This is numerically a very poor approach because of the singularity of the weight function at $x = \pm 1$. The result will be very sensitive to the actual sampling near the outer boundaries, leading to an ill-conditioned problem. One way out is to by-pass the weighting by using a non-uniform sampling with more data points concentrated near the boundaries than at the centre of the interval, so that the number of points in the interval dx is proportional to,

$$\frac{dx}{(1 - x^2)^{1/2}} \tag{1.23}$$

This is equivalent to using the variable $\theta = \arccos x$ in (1.19), leading to

$$c_j = \frac{2}{\pi} \int_0^\pi f(\cos\theta) T_j(\cos\theta)\, d\theta \tag{1.24}$$

Therefore a stable solution to the discrete computation of the Chebyshev expansion is achievable with an unweighed least squares fit with a uniform sampling

$$\theta_1, \theta_2, \cdots, \theta_n, \qquad 0 \leq \theta_i \leq \pi \tag{1.25}$$

in θ. Using 1000 points uniformly distributed one gets,

$$q_3(x) \approx 2.5698\, x - 2.6673\, x^3 \tag{1.26}$$

instead of (1.21). Again the two would get closer to each other with a denser sampling in θ. One sees that the evaluation can be computationally demanding to achieve a good accuracy. This assumes that the function $f(x)$ can be evaluated accurately at any point x in the interval $[-1, 1]$, which is not necessarily true when the data points come from a numerical integration, though this can normally be done to any prescribed accuracy with a standard polynomial interpolation between two entries.

1.5 Best approximation

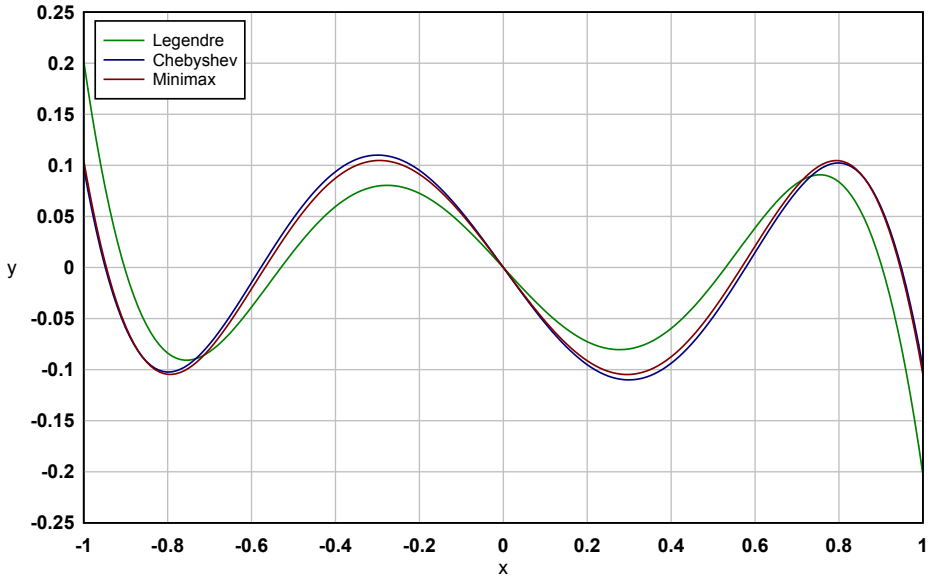


FIGURE 1.4: Comparison of the errors for three approximations of $\sin(\pi x)$ with a polynomial of third degree when $-1 \leq x \leq 1$. The plot gives the error for a Legendre (ie unweighed polynomial fit) (green), Chebyshev (blue) and minimax approximations. It shows the near identity between the last two methods and the growth of the error in Legendre expansion close to the outer boundaries.

To conclude this introductory section I consider now the general problem of finding a polynomial approximation of degree n of $f(x)$, with $f \in \mathcal{C}[a,b]$ which produces the least error on the interval $[a,b]$. We then search the polynomial $q_n^*(x)$ of $\mathcal{P}_n$ such that

$$||f - q_n^*||_\infty \leq ||f - q_n||_\infty \quad \text{for any} \quad q_n \in \mathcal{P}_n \tag{1.27}$$

In other words, there does not exist a polynomial of degree $\leq n$ that can approximate $f(x)$ with a smaller maximum error than $q_n^*(x)$. In practice this best approximation is much harder to obtain that any of the polynomial approximations shown earlier in this section, although efficient algorithms exist. Using the Remes or exchange algorithm (see Rivlin (1981) or K. Atkinson (1989)) implemented in Maple, one gets the best polynomial approximation of $\sin(\pi x)$ by a polynomial of third degree as,

$$q_3^*(x) = 2.559018\,x - 2.69488\,x^3 \tag{1.28}$$

and

$$||f - q_3^*||_\infty = 0.105 \tag{1.29a}$$

$$||f - q_3^*||_2 = 0.104 \tag{1.29b}$$

The best approximation (1.28) is very similar to (1.21) and this fact is a very general property discussed in Sect. 2. Fig. 1.4 shows a comparison between the errors for the Legendre (same as uniform least squares), Chebyshev and Minimax approximations for $\sin \pi x, \quad -1 \leq x \leq 1$.

2 Chebyshev representation

This section summarises the most important definitions and properties of the standard Chebyshev expansion. It explains also why this expansion is so attractive to approximate a function that could be complex or too much CPU demanding to evaluate otherwise. The presentation deals primarily with the theoretical properties that apply to general functions. This section is an update of a note of 2003 written while preparing a version of the Gaia and Earth ephemeris for the Gaia simulation (GAIA-FM-013).

2.1 Basic definition

The expansion of a function $f(x)$ in a series of Chebyshev polynomials is given by,

$$\tilde{f}(x) = \frac{1}{2}c_0 + \sum_{k=1}^{\infty} c_k\, T_k(x) \tag{2.1}$$

where x lies in the interval $[-1,\ 1]$ and $T_k(x)$ is the kth Chebyshev polynomial of first kind.

The c_k are so chosen as to minimise,

$$\min_{c_0,\cdots,c_k,\cdots} \int_{-1}^{1} [\tilde{f}(x) - f(x)]^2 w(x) dx \tag{2.2}$$

where $w(x) = (1-x^2)^{-1/2}$. Naturally condition (2.2) is the key statement, while the representation (2.1) is just the most convenient with the weight $w(x)$. Any other representation but using (2.2) to compute the coefficients will lead to the same approximation, but expanded on different basis functions.

It is customary in Fourier-like expansions to rewrite (2.1) as,

$$f(x) = \sum_{k=0}^{\infty}{}' c_k\, T_k(x) \tag{2.3}$$

the prime indicating that the first term of the sum must be halved (one could say that it would be better to provide $a_0/2$ for the first term, but the interest of this notation is to have a single general expression to compute the a_i, with no exception for a_0.). For a continuous, square integrable and piecewise differentiable function the expansion (2.1) is uniformly convergent. The approximation of $f(x)$ by a truncated expansion of degree m is then given by,

$$f(x) \approx \phi(x) = \sum_{k=0}^{m}{}' c_k\, T_k(x) \tag{2.4}$$

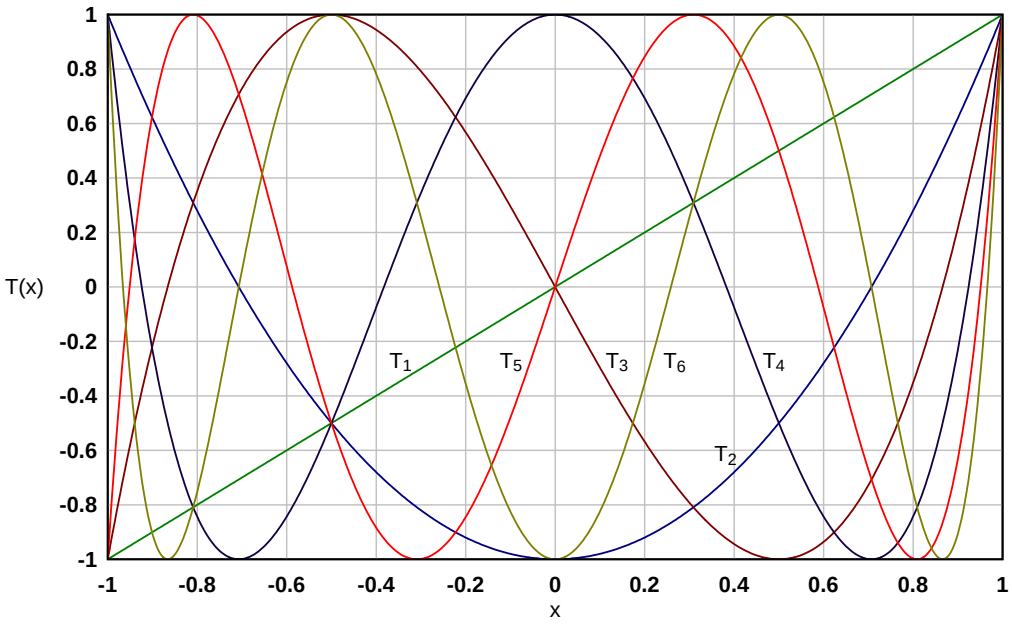


FIGURE 2.1: Plots of the first Chebyshev polynomials of the first kind for degree 1 to 6. A polynomial of degree k has k zeros and $k+1$ extremums equal to ± 1 in $[-1,\ 1]$.

2.2 Chebyshev polynomials of first kind

The Chebyshev polynomial (of first kind) of degree k is given by,

$$T_k(x) = \cos(k \arccos x) \tag{2.5}$$

and satisfies the recurrence equation [1],

$$T_{k+1}(x) = 2x\,T_k(x) - T_{k-1}(x) \tag{2.6}$$

with

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \end{aligned} \tag{2.7}$$

[1] It was very likely used by the Renaissance mathematician F. Viete (1540-1603) to provide a solution to the challenge sent in 1593 to the European best mathematicians of that time by Adriaan van Roomen (latinised as Adrianus Romanus). The problem required to find a numerical solution to a polynomial equation of degree 45. It was solved overnight by Viète who realised that there was an underlying trigonometric relation equivalent to the expression of $\sin(45\,x)$ in a polynomial in $\sin x$.

The evaluation of $\phi(x)$ from the c_k can be done with comparable efficiency with (2.6) or with the today more popular Clenshaw downward recurrence formula, as indicated in *Numerical Recipe* Press et al. (1992).

Property 1. *The leading coefficient of* $T_k(x), k \geq 1$ *(coefficient of* x^k*) is* $a_k^{(k)} = 2^{k-1}$

Proof. If

$$T_k(x) = \sum_{i=0}^{k} a_i^{(k)} x^i$$

then it is easy to derive from (2.6) that $a_{k+1}^{(k+1)} = 2a_k^{(k)}$. With $a_1^{(1)} = 1$, this leads to the result. □

A standard notation for the polynomial normalized with unit leading coefficient is then $\widetilde{T}_k(x) = T_k(x)/2^{k-1}$. It is often useful to introduce $x = \cos\theta,\ \theta$ in $[0, \pi]$ and write,

$$T_k(\cos\theta) = \cos k\theta \tag{2.8}$$

easily derived from the recurrence equation.

From this equivalence one gets the zeros $\xi_1,\ \xi_2, \cdots, \xi_k$ of $T_k(x)$ as,

$$\xi_j = \cos\left(\frac{(2j-1)\pi}{2k}\right), \quad j = 1, 2, \cdots, k \tag{2.9}$$

and the $k+1$ abscissas of the upper and lower bounds,

$$\zeta_j = \cos\left(\frac{j\pi}{k}\right), \quad j = 0, 1, \cdots, k \tag{2.10}$$

Property 2. *The Chebyshev polynomials are orthogonal to each other on the interval* $[-1,\ 1]$ *with the weight function* $w(x) = (1-x^2)^{-1/2}$ *so that,*

$$\int_{-1}^{1} \frac{T_m(x)\, T_n(x)}{(1-x^2)^{1/2}}\, dx = \int_0^{\pi} \cos m\theta\, \cos n\theta\, d\theta = \begin{cases} 0 & \text{if} \quad m \neq n \\ \pi & \text{if} \quad m = n = 0 \\ \pi/2 & \text{if} \quad m = n \neq 0 \end{cases} \tag{2.11}$$

As a result of the orthogonality we can write the Fourier coefficients (in the least squares sense) of the expansion of a function $f(x)$ in a series of Chebyshev polynomials (Eq. 2.1) as,

$$c_j = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x)T_j(x)}{(1-x^2)^{1/2}}\, dx, \quad j = 0,\, 1,\, \cdots \tag{2.12}$$

More generally for a function $f(y)$ over the interval $[a,\, b]$ the affine transformation,

$$x = \frac{y-(a+b)/2}{(b-a)/2} \tag{2.13}$$

$$y = \frac{(b-a)}{2}\, x + \frac{(a+b)}{2} \tag{2.14}$$

maps the interval $[a,\, b]$ of y on the interval $[-1,\, 1]$ of x.

The Chebyshev polynomials are so valuable in representing functions over a given interval because of the following property :

Property 3. *Of all polynomials of degree k with leading coefficient equal to 1, the Chebyshev polynomial of degree k divided by 2^{k-1} oscillates with minimum maximum amplitude on the interval $[-1,\, 1]$.*

Proof. The proof is by contradiction and is very elegant. We assume that there exists another polynomial $Q_k(x)$ with leading coefficient 1 which has smaller maximum amplitude on $[-1,\, 1]$. Then the polynomial $P_{k-1}(x) = \widetilde{T}_k(x) - Q_k(x)$ is of degree $k-1$. At each of the $k+1$ extremum ζ_i, $i = 0.1 \cdots, k$ of $\widetilde{T}_k(x)$ the following inequality holds : $|Q_k(x)| < |\widetilde{T}_k(x)|$. Therefore $P_{k-1}(x)$ alternates in sign from one extremum to the next, which means that $P_{k-1}(x)$ has k zeros, an impossibility for a polynomial of degree $k-1$ and this contradicts the initial hypothesis. □

This means that among all the polynomials $P_k(x)$ of degree k and with $a_k = 1$, the Chebyshev polynomial $T_k(x)/2^{k-1}$ is the closest to the $x-$axis. So

$$2^{-(k-1)} = \max_{x\in[-1,\,1]} |\widetilde{T_k}(x)| \leq \max_{x\in[-1,\,1]} |P_k(x)| \text{ for any } P_k \in \pi_k \tag{2.15}$$

More generally one calls the best approximation of $f(x)$ on an interval $[a, b]$ by a polynomial p_k of degree k the polynomial p_k^* such that,

$$\|f - p_k^*\|_\infty = \min_{p\in\pi_k} \|f - p_k\|_\infty \tag{2.16}$$

In other words the particular polynomial p_k^* is as close to the function $f(x)$ as it is possible for polynomials of degree k. It happens that best approximations (called also minimax approximations or uniform approximations) are very difficult to construct, but near minimax approximations are quite often obtained with a Chebyshev expansion truncated at degree k. In particular, as a consequence of Prop. 3 one has the following result :

Property 4. *The best approximation p_k^* of x^n by a polynomial of degree $k < n$ is given by $x^n - T_n/2^{n-1}$.*

Then the remainder $x^n - p_k^*$ is proportional to $T_n(x)$ and vanishes at the zeros (Eq. 2.9) of $T_n(x)$. Therefore $p_k^*(x)$ can be obtained by interpolating x^n at $\xi_1, \xi_2, \cdots, \xi_n$ (hence $k \leq n-1$) where $T_n(\xi_j) = 0$.

Since for any function $f(x)$ the best approximation p_n^* is such that the error must alternate $n+1$ times between equal (in magnitude) upper and lower bounds, there exists also $n+1$ points η_i, $i = 1, \cdots, n+1$ at which the error vanishes, that is to say at which p_n^* interpolates $f(x)$. One then could in principle construct the best approximation by interpolation if one knew where to interpolate. With Prop. 4 it is natural to consider the Chebyshev points as a reasonable all-purpose choice. In practice such an interpolation is known to give approximations which are at the same time very close to the best approximation and to the Chebyshev expansion in the least squares sense (Eqs. 2.1 and 2.12), relieving us from the burden of evaluating integrals.

2.3 Chebyshev polynomials of second kind

The Chebyshev polynomials of second kind are less important for the approximation theory, but are used later in this note and this is the right place to introduce them.

The Chebyshev polynomial of second kind and of degree k is given by,

$$U_k(x) = \frac{\sin((k+1)\theta)}{\sin\theta} \tag{2.17}$$

where $x = \cos\theta$. It is quickly checked that (2.17) is a polynomial of degree k in $\cos\theta$. These polynomials satisfy the same recurrence recurrence equation as those of the first kind, namely,

$$U_{k+1}(x) = 2x\,U_k(x) - U_{k-1}(x) \tag{2.18}$$

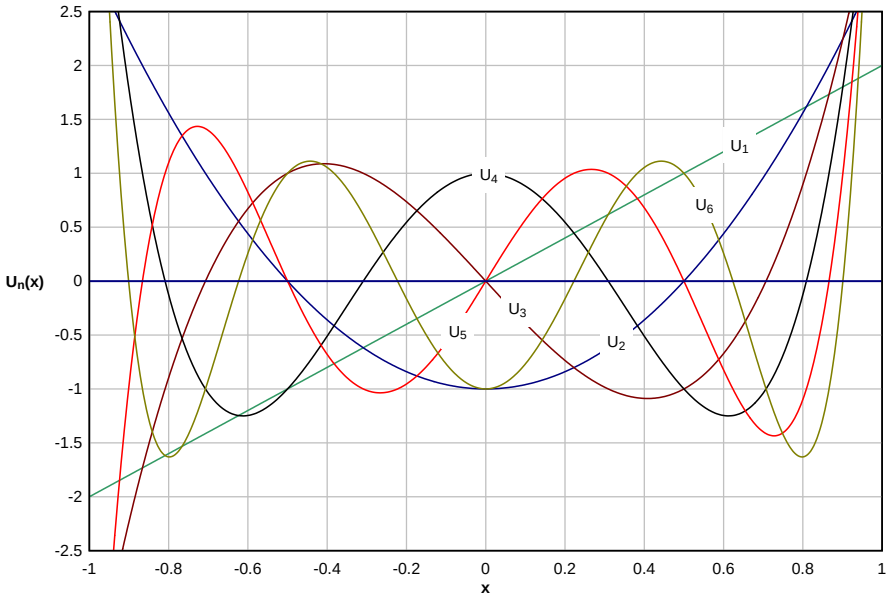


FIGURE 2.2: Plots of the first Chebyshev polynomials of the second kind for degree 1 to 6. A polynomial of degree k has k zeros and $k+1$ extremums in $[-1,\ 1]$.

but with the starting values,

$$\begin{aligned} U_0(x) &= 1 \\ U_1(x) &= 2x \end{aligned} \tag{2.19}$$

instead of (2.7). The first few polynomials are plotted in Fig. 2.2.

They are related to the $T_k(x)$ (for $k > 0$) by,

$$\frac{dT_k}{dx} = k\frac{\sin k\theta}{\sin\theta} = kU_{k-1}(x) \tag{2.20}$$

allowing to evaluate quickly the first derivative of the expansion (2.1) as,

$$\frac{d\tilde{f}(x)}{dx} = \sum_{k=1}^{\infty} kc_k\, U_{k-1}(x) \tag{2.21}$$

Hence no additional effort is needed to evaluate the first derivative of the expansion. Whether the truncated version to degree m of (2.21) is a good expansion of $f'(x)$ is another story (the

answer is: in general this is not and it is much preferable to expand the derivative directly, as shown in Sect. 4.2 of (GAIA-FM-013)). The convergence of (2.21) is not guaranteed, when the c_k do not decrease faster than $1/k$. Not surprisingly, they do when $f(x) \in \mathcal{C}^1[-1,1]$, that is to say when $f(x)$ has a continuous first derivative, or when trying to approximate the derivative makes sense.

There are two others useful links between the polynomials of first and second kind,

$$2T_k(x) = U_k(x) - U_{k-2}(x) \tag{2.22a}$$

$$T_k(x) = U_k(x) - xU_{k-1}(x) \tag{2.22b}$$

which can be derived with trigonometrical identities.

The polynomial $U_k(x)$ has k zeros in the interval $[-1,1]$, given by,

$$\tilde{\xi}_j = \cos\left(\frac{j\pi}{k+1}\right), \quad j = 1, 2, \cdots, k \tag{2.23}$$

or with (2.10), they are identical to the inner extremums of $T_{k+1}(x)$.

Finally, the set $\{U_0(x), U_1(x), \cdots, U_k(x), \cdots\}$ constitutes a sequence of orthogonal polynomials. More precisely one has,

Property 5. *The Chebyshev polynomials of second kind are orthogonal to each other on the interval* $[-1,\ 1]$ *with the weight function* $w(x) = (1-x^2)^{1/2}$ *so that,*

$$\int_{-1}^{1} U_m(x)\, U_n(x)\, (1-x^2)^{1/2}\, dx = \begin{cases} 0 & \text{if} \quad m \neq n \\ \pi/2 & \text{if} \quad m = n \end{cases} \tag{2.24}$$

2.4 Explicit forms

The first few Chebyshev polynomials of first kind are,

$$
\begin{aligned}
T_0(x) &= 1 \\
T_1(x) &= x \\
T_2(x) &= 2x^2 - 1 \\
T_3(x) &= 4x^3 - 3x \\
T_4(x) &= 8x^4 - 8x^2 + 1 \\
T_5(x) &= 16x^5 - 20x^3 + 5x \\
T_6(x) &= 32x^6 - 48x^4 + 18x^2 - 1
\end{aligned}
\tag{2.25}
$$

The first few Chebyshev polynomials of second kind are,

$$
\begin{aligned}
U_0(x) &= 1 \\
U_1(x) &= 2x \\
U_2(x) &= 4x^2 - 1 \\
U_3(x) &= 8x^3 - 4x \\
U_4(x) &= 16x^4 - 12x^2 + 1 \\
U_5(x) &= 32x^5 - 32x^3 + 6x \\
U_6(x) &= 64x^6 - 80x^4 + 24x^2 - 1
\end{aligned}
\tag{2.26}
$$

One can readily check in Eqs. 2.25–2.26, the relations (2.20)–(2.22) between the two kinds of polynomials.

3 Accuracy of a Chebyshev expansion

In this section I look for some guiding rules to evaluate the error made by truncating the Chebyshev approximation to a finite degree, in particular when the approximated function is quasiperiodic, the most relevant situation met in solar system dynamics. The trade-off between the degree of the expansion and the size of the sub-intervals to reach a given accuracy is also discussed.

3.1 General case

Unlike the interpolation theory, there is no general rule to bound the error done in the approximation of a function $f(x) \in \mathcal{C}[-1,1]$ by a truncated Chebyshev expansion, that is to say using the approximation $\phi_n(x) \approx f(x)$ with,

$$\phi_n(x) = {\sum_{k=0}^{k=n}}' c_k\, T_k(x) \tag{3.1}$$

and with

$$c_k = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x)\, T_k(x)}{\sqrt{1-x^2}}\, dx \tag{3.2}$$

Without any additional assumption on $f(x)$, we can at least state that,

$$|f(x) - \phi_n(x)| = \Big| \sum_{k=n+1}^{\infty} c_k T_k(x) \Big| \leq \sum_{k=n+1}^{\infty} |c_k| \tag{3.3}$$

and if the c_k decreases fast enough with k the above series will converge. In addition, if the c_k alternate in sign, therefore $|c_{n+1}|$ bounds the error. There is a more general result telling that if $f(x) \in \mathcal{C}_2[-1,1]$, then there exists a constant a, depending only on f and not on k, such that,

$$|c_k| \leq \frac{a}{k^2} \tag{3.4}$$

and the Chebyshev expansion has a pointwise uniform convergence, with

$$\lim_{n\to\infty} ||f - \phi_n|| = 0 \tag{3.5}$$

Finally the following property is more useful for practical applications.

Property 6. *For any function $f(x) \in \mathcal{C}_k[\alpha,\beta]$, one has*

$$|c_k| \leq \frac{\max|f^k(x)|}{2^{k-1}k!}\,\frac{(\beta-\alpha)^k}{2^k} \tag{3.6}$$

Proof. Without loss of generality one can take $\alpha = -1$ and $\beta = 1$, and restore the original variable with an affine transformation. With the continuity of the k first derivatives on $[-1,1]$ and with $-1 \leq x \leq 1$ we have the Taylor expansion around $x_0 = 0$ as,

$$f(x) = f(0) + \frac{f'(0)}{1!}x + \frac{f^2(0)}{2!}x^2 + \cdots + \frac{f^{k-1}(0)}{(k-1)!}x^{k-1} + \frac{f^k(\tau)}{k!}x^k \tag{3.7}$$

where $\tau \in [-1,1]$. Now in a Chebyshev expansion of $f(x)$ one has,

$$c_k = \int_{-1}^{1} \frac{f(x)\,T_k(x)}{(1-x^2)^{1/2}}dx \tag{3.8}$$

and since,

$$\int_{-1}^{1} \frac{x^j\,T_k(x)}{(1-x^2)^{1/2}}dx = \begin{cases} 0, & \text{if } j < k \\ \frac{1}{2^{k-1}} & \text{if } j = k \end{cases} \tag{3.9}$$

one obtains,

$$c_k = \frac{f^k(\tau)}{2^{k-1}\,k!} \tag{3.10}$$

and

$$|c_k| \leq \frac{\max_{x\in[-1,1]}|f^k(x)|}{2^{k-1}\,k!}, \quad -1 \leq x \leq 1 \tag{3.11}$$

Returning to the original interval $[\alpha,\beta]$ gives (3.6). □

Eq. 3.11 is the basis of a widely used criterion to determine the truncation error when stopping the Chebyshev expansion to degree n. For most functions of practical interest the magnitude of the c_k decreases relatively quickly with increasing k, and even if they do not alternate in sign, it is generally accepted that the truncation error is well estimated by,

$$|f(x) - \phi_n(x)| \sim c_{n+1}\,T_{n+1}(x) \tag{3.12}$$

With $|T_n(x)| \leq 1$, the maximum error is then of the order, if not bounded by, $|c_{n+1}|$.

3.2 Quasi periodic functions

For Gaia, and more generally in ephemeris representations, the functions $f(t)$ of interest are generally the linear combination of few secular terms expressed in polynomial form and a quasi-periodic component, expanded over trigonometric functions. When expanded over an interval Δt around a reference time t_0, say between $[t_0 - \Delta t/2, t_0 + \Delta t/2]$, the polynomial secular terms are projected on a *finite* number of Chebyshev polynomials and are not a source of worry. I concentrate then on the trigonometric terms of the kind

$$\cos(2\pi \frac{t}{T}) \tag{3.13}$$

where T is the period.

One may ask how to chose the interval Δt and the smallest degree n to ensure a given accuracy in the expansion. It is natural to expect that for $\Delta t/T \geq 1$, the degree will have to be large, because of the numerous oscillations of the function over the time interval. We can play both on the interval and the degree to achieve the final accuracy. Increasing the degree will cost more to make one evaluation. Decreasing Δt will increase the number of sub-intervals and the total amount of data storage. Knowing that the representation is needed over a total duration D much larger than Δt, the tradeoff is between the volume of the representation proportional to $D/\Delta t\, n$ and the computing cost for an evaluation directly proportional to the degree n, and the accuracy $\epsilon(n, \Delta t/T)$.

In fact, with this simple function, one can evaluate the Chebyshev coefficients analytically and estimate the error when truncating the expansion. Take the slightly more general function,

$$\cos(2\pi \frac{t}{T} - \psi) = \cos(\pi \frac{\Delta t}{T}\, x - \psi) \tag{3.14}$$

with a phase ψ and $x = 2t/\Delta t$, with x in $(-1, 1)$.

Then one has,

$$\begin{aligned} c_k &= \frac{2}{\pi} \int_{-1}^{1} \cos\left(\pi \frac{\Delta t}{T}\, x - \psi\right) \frac{T_k(x)}{\sqrt{1-x^2}}\, dx \\ &= \frac{2}{\pi} \int_{0}^{\pi} \cos\left(\pi \frac{\Delta t}{T} \cos\theta - \psi\right) \cos k\theta\, d\theta \end{aligned} \tag{3.15}$$

Using now the well known integral form for the Bessel functions,

$$J_n(\alpha) = \frac{1}{\pi}\int_0^{\pi} \cos(n\theta - \alpha\sin\theta)\, d\theta \tag{3.16}$$

one has also,

$$c_k = \begin{cases} (-1)^p\, 2\cos\psi\, J_k\left(\pi\frac{\Delta t}{T}\right) & \text{if } k = 2p \\ (-1)^p\, 2\sin\psi\, J_k\left(\pi\frac{\Delta t}{T}\right) & \text{if } k = 2p+1 \end{cases} \tag{3.17}$$

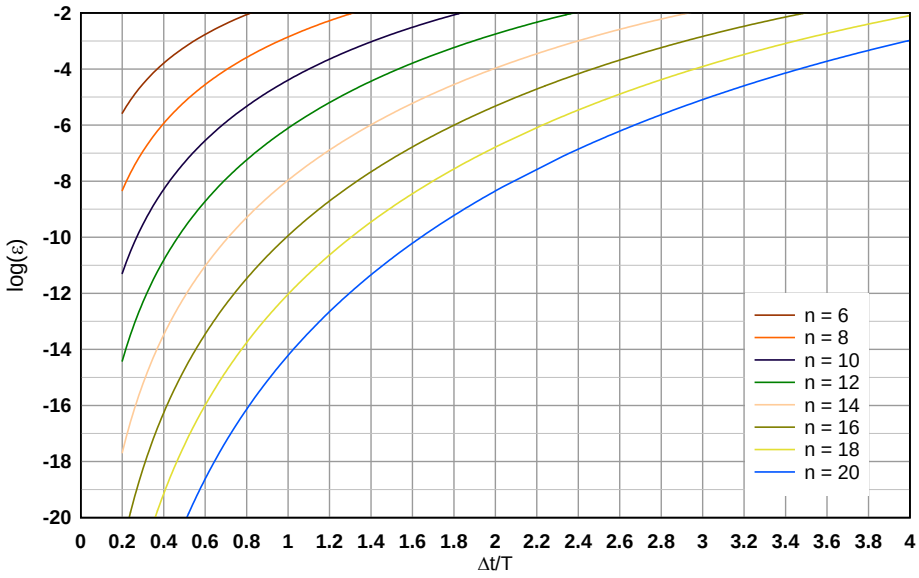


FIGURE 3.1: Amplitude of the nth coefficient in the Chebyshev expansion of $\cos(\pi\Delta t/T\, x)$, with x in $(-1, 1)$, as a function of the window size Δt expressed in period. This is also a good estimate of the error for a truncated expansion to degree $n-1$.

giving essentially,

$$|c_k| \approx 2J_k\left(\pi\frac{\Delta t}{T}\right) \tag{3.18}$$

and with the asymptotic expression of the Bessel function,

$$J_k(x) \sim \frac{1}{n!}\left(\frac{x}{2}\right)^k, \quad x \ll \sqrt{k+1} \tag{3.19}$$

one has also with $n! \sim \sqrt{2\pi n}(n/e)^n$

$$\boxed{|c_k| \sim \sqrt{\frac{2}{k\pi}} \left[\frac{e\pi}{2k}\frac{\Delta t}{T}\right]^k} \tag{3.20}$$

This gives the magnitude of the kth coefficients in the Chebyshev expansion of a periodic function of unit amplitude, as a function of the size of the time interval, scaled by the period. We see, as expected, that for $\Delta t < T$, the Chebyshev coefficients decrease quickly with the order of the polynomial. More precisely one has for this case,

$$\frac{|c_{k+1}|}{|c_k|} = \sqrt{\frac{k}{k+1}}\,\frac{k^k}{(k+1)^{k+1}} < \frac{k^k}{(k+1)^{k+1}} < \frac{1}{k} \tag{3.21}$$

which shows that as soon as the degree is larger than 10, the error in the Chebyshev expansion decreases by more than one order of magnitude for each increase of one unit.

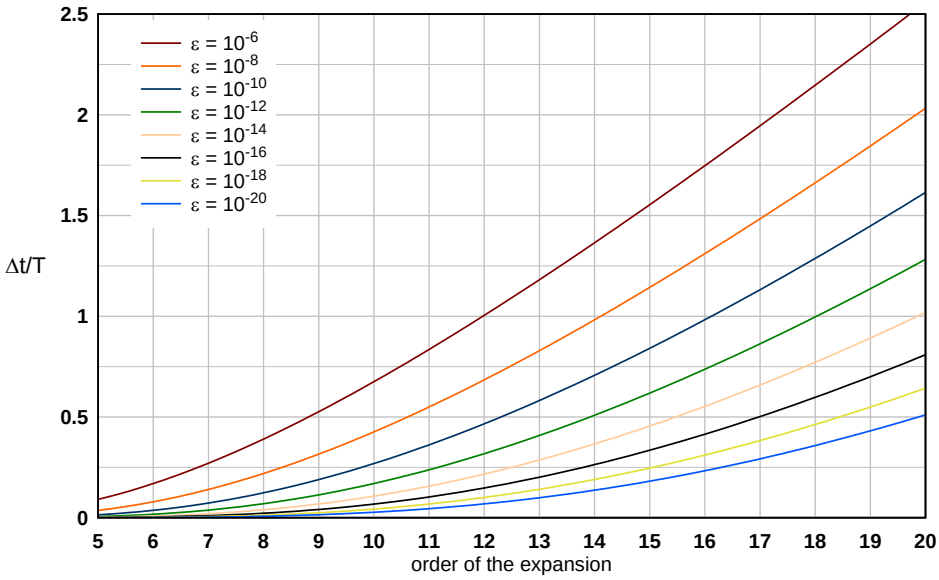


FIGURE 3.2: Locus in the plane n-$\Delta t/T$ of the curves of constant accuracy in a Chebyshev expansion of $\cos(\pi \Delta t/T\, x)$, with x in $(-1, 1)$. For a preset accuracy, one can play between the order n of the expansion and the size of time interval Δt.

The coefficients (in fact computed with the exact expression using Bessel functions) are plotted

in Fig. 3.1, giving the magnitude of the coefficients as a function of Δt and for different orders of the expansion. The coefficients fall off rather quickly as soon as they are less than $\sim 10^{-5}$ and the first neglected coefficients provided a good estimate of the truncation error made by replacing the infinite expansion by a truncated version. In practice, the useful question is rather finding the degree for a prescribed maximum error. This is shown in Fig. 3.2, giving the locus of constant error in the plane $n, \Delta t/T$. The curve are drawn continuous, but obviously only integral value of the order are of interest here.

4 Computation of the Chebyshev coefficients

4.1 Principles

In the least squares sense the computation of the Chebyshev coefficients reduces to the evaluation of the integrals,

$$c_j = \frac{2}{\pi} \int_{-1}^{1} \frac{f(x)T_j(x)}{(1-x^2)^{1/2}}\, dx, \quad j = 0,\, 1,\, \cdots \tag{4.1}$$

Except in very specific cases with elementary functions, these coefficients must be computed with numerical methods, described in Numerical Analysis textbooks (K. Atkinson (1989), Ralston et al. (1978)). Given the singularity of the integrand with the $1/(1-x^2)^{1/2}$ weights, any among the Newton-Cotes formulas using the end-points would fail. It happens that the Gauss-Chebyshev numerical integration scheme is designed to handle this kind of singularity. Fortunately it is also very easy to implement to any degree, unlike the Gauss method which requires special tables to find the integration nodes and the coefficients. The scheme is as follows at the order n,

$$\int_{-1}^{1} \frac{f(x)dx}{\sqrt{1-x^2}} \approx \sum_{i=1}^{n} w^{(n)} f(\xi_i^{(n)}) \tag{4.2}$$

where

$$w^{(n)} = \pi/n \tag{4.3}$$

(depends only on the order and not on i) and the nodes $\xi_i^{(n)}$ are the roots $T_n(x)$ given by,

$$\xi_i^{(n)} = \cos\left(\frac{(2i-1)\pi}{2n}\right) \tag{4.4}$$

such that

$$T_n(\xi_i^{(n)}) = 0, \;\; i = 1, 2, \cdots, n. \tag{4.5}$$

The integration with (4.2) is exact for polynomials up to degree $2n-1$. Using the Gauss-

Chebyshev integration scheme to degree n, on can then evaluate the c_j in (4.1) by,

$$C_j = \frac{2}{n} \sum_{i=1}^{n} f(\xi_i)\, T_j(\xi_i), \quad j = 0,\, 1,\, \cdots,\, k \tag{4.6}$$

where the order n must be chosen to ensure $C_j \approx c_j$. There is no general rule to achieve this choice, which depends on the degree of differentiability of $f(x)$.

4.2 Relationship to interpolation

The error in the truncated Chebyshev approximation of degree n is generally dominated by the first neglected polynomial $T_{n+1}(x)$, although in some peculiar situations it could be even a polynomial of higher degree (eg: an odd or even function may have only odd or even non-zero coefficients). Consider the standard case with $c_{n+1} \neq 0$. Since $T_{n+1}(x)$ has $n+1$ zeros, the truncated expansion takes the same value as the function at $n+1$ points in the interval $(-1, 1)$. If we knew these points in advance, one could get exactly the same approximation by interpolating the function at these points with a polynomial of degree n to find the $n+1$ coefficients $c_0, \cdots, c_n$. One of the interests in taking this view for the approximation is the fact the Lagrange interpolation has an handy expression for the remainder as a function of the sampling and some other terms independent of this sampling.

Quite generally for the interpolation of a function $f(y)$ at the k distinct points $y_1,\, y_2,\, \cdots,\, y_k$ by a polynomial $P_{k-1}(y)$ it is known that the there exists $\zeta(y) \in [y_1,\, y_k]$ such that,

$$E(y) = f(y) - P_{k-1}(y) = \frac{f^k(\zeta)}{k!}\,(y - y_1)(y - y_2)\cdots(y - y_k) \tag{4.7}$$

In view of comparing errors on different intervals it is interesting to use the normalised variable $x \in [-1,\, 1]$ and for an interpolation between $y_1 = \alpha$ and $y_k = \beta$ one has instead,

$$E(x) = \frac{f^k(\zeta)}{k!}\,\frac{(\beta - \alpha)^k}{2^k}\,(x - x_1)(x - x_2)\cdots(x - x_k) \tag{4.8}$$

In particular one has $x_i = \cos\left(\frac{(2i-1)\pi}{2k}\right), i = 1, \cdots, k$ for an interpolation of degree $k-1$ at the Chebyshev points, $x_i = \cos\left(\frac{i\pi}{(k-1)}\right),\ i = 0, \cdots, k-1$ for the Lobatto [2] points (max of

[2] There is no general terminology to refer to these points. Given the close relationship between the interpolation using end-points and the Lobatto numerical quadrature, it is relevant to name them after him. Rehuel Lobatto (1797-1866) was a Dutch mathematician noted for his contributions to statistics and their application to life insurance.

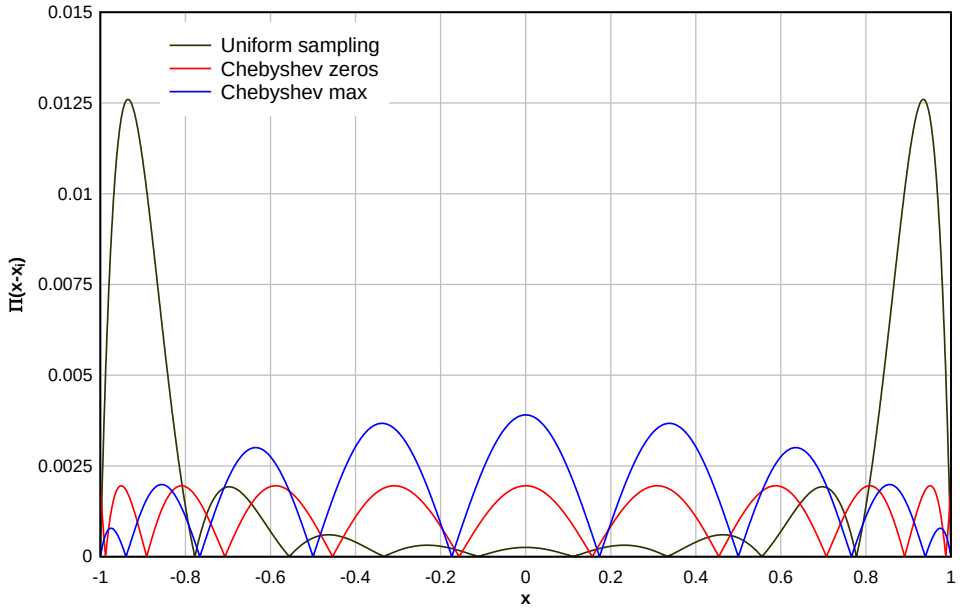


FIGURE 4.1: Comparison of the main adjustable factor in the remainder of an interpolation according to the choice of the nodes. In the plot a degree 9 is common to the three samplings, meaning 10 interpolation nodes. The black curve corresponds to a uniform sampling, the blue one for a sampling at the Chebyshev points (the zeros of $T_n(x)$) and the red curve is obtained for a sampling at the maximum of $T_{n-1}(x)$ and the two end points (Lobatto points).

the Chebyshev polynomials and the two end points) and $x_i = -1 + \frac{2(i-1)}{k-1}, i = 1, \cdots, k$ for a regular distribution of the interpolation points. The error term (4.8) is bounded by the max of the kth derivative in the interval $(-1, 1)$ and that of a polynomial fully determined by the sampling. The comparison of this latter factor for the three samplings (regular, Chebyshev,) is shown in Fig. 4.1 for 10 interpolation nodes. The difference between the three cases is striking and one sees several interesting features:

- Runge phenomena at the boundaries of the uniform interpolation. It grows with the number of nodes preventing a uniform convergence of the interpolation over the interval and generating a strong oscillatory behaviour between the nodes.
- Regular distribution of the interpolation errors when the Chebyshev points are used, with a higher density near the boundaries than at the centre. The maximums are all identical to $1/2^9$, the smallest maximum value of a polynomial of degree 9 and leading coefficient equal to 1. The error is non-zero at the boundaries.
- A good compromise is obtained with the interpolation at the Lobatto points (in fact the zeros of $(1 - x^2)U_{n-1}(x)$ where $U_n(x)$ is the Chebyshev polynomial of second kind). The error is fairly regular and the interpolation nodes include the end points.

Therefore it appears quite appealing to investigate more closely how one can obtain these two interpolations directly in the form of a Chebyshev expansion. This is the topic of the next two subsections where a practical method is given to build the interpolation at the Chebyshev and Lobatto nodes and express the results in the form of a Chebyshev-like expansion.

Remark: Given the extremely good accuracy of the uniform interpolation (equally spaced nodes) in the central part of the total interval, it is tempting to restrict its use to typically half the interval around the origin. Then, the interpolating polynomial would be constructed over overlapping intervals of size 2Δ centred at x_i, but used only over $[x_i - \Delta/2, x_i + \Delta/2]$, with $x_{i+1} = x_i + \Delta$. One then would benefit from the excellent accuracy over half of the interpolating interval as shown in Fig. 4.1, typically better than the error on the Chebyshev expansion built on the interval 2Δ. This is true, but the comparison is somewhat misleading: for a total representation of length T, one would use $T/(2\Delta)$ sub-intervals in the Chebyshev expansions, and twice as many with the uniform sampling. Therefore the relevant comparison must be done with a Chebyshev expansion built also on intervals of size Δ, non overlapping. We know from (3.6), that dividing the size of the sub-interval by two, decreases the error for an expansion to degree k by 2^k, a much better gain than the one obtained by simply using the central part of the Lagrange interpolation with a regular sampling. Thus what looked at first glance like a bright idea does not resist a closer examination. This comment is based on an earlier discussion with R. Cayrel, an astronomer with the Observatory of Paris.

4.3 Chebyshev approximation by interpolation

For a given function $f(x)$ the interpolation at the Chebyshev points could be in principle obtained by a Lagrange interpolation using for example the divided difference scheme or the evaluation procedure based on the Neuville algorithm, both methods being explained in any numerical analysis textbook. But publishing divided differences instead of Chebyshev coefficients is not so convenient as no truncation is permitted if one wishes to use the full interval upon which the interpolation is built.

Fortunately there is an important property that renders the computation of the interpolation on Chebyshev points very simple, yielding directly the interpolating polynomial projected on the Chebyshev basis.

Property 7. *If the coefficients C_0, C_1, $\cdots$, C_k of a Chebyshev expansion are computed as :*

$$C_j = \frac{2}{k+1} \sum_{i=1}^{k+1} f(\xi_i)\, T_j(\xi_i), \quad j = 0,\, 1,\, \cdots,\, k \tag{4.9}$$

where $T_{k+1}(\xi_i) = 0$, that is to say the ξ_i $(i = 1, \cdots, k+1)$ are the zeros of $T_{k+1}(x)$, the resulting polynomial $\Phi(x) = \sum'^{k}_{j=0} C_j\, T_j(x)$ of degree k interpolates $f(x)$ at the $k+1$ Chebyshev points.

Remark: this means that the integrals (4.1) up to $j = k$ are numerically computed with a Gauss-Chebyshev scheme of order $k + 1$.

Proof. Let

$$\Phi(x) = \sum_{j=0}^{j=k}{}' C_j\, T_j(x)$$

be a Chebyshev representation of $f(x),\ x \in [-1,\, 1]$, with

$$C_j = \frac{2}{k+1} \sum_{i=1}^{k+1} f(\xi_i)\, T_j(\xi_i), \quad j = 0,\, 1,\, \cdots,\, k \tag{4.10}$$

where $T_{k+1}(\xi_i) = 0$. Then,

$$\Phi(\xi_r) = \sum_{j=0}^{j=k}{}' C_j\, T_j(\xi_r), \quad r = 1, \cdots, k+1 \tag{4.11}$$

and by inserting (4.10) in (4.11) we have,

$$\Phi(\xi_r) = \frac{2}{k+1} \sum_{j=0}^{j=k}{}' \sum_{i=1}^{k+1} f(\xi_i)\, T_j(\xi_i)\, T_j(\xi_r),, \quad r = 1,\, \cdots,\, k+1$$

Using the definition of $T_k(x)$ and the values of the roots ξ_r of $T_{k+1}(x)$ the right hand side becomes,

$$\frac{2}{k+1} \sum_{i=1}^{k+1} f(\xi_i) \sum_{j=0}^{k}{}' \cos\left(j \frac{(2i-1)\pi}{2(k+1)}\right) \cos\left(j \frac{(2r-1)\pi}{2(k+1)}\right) \tag{4.12}$$

Transforming now the product of cosines into a sum and using the classical results,

$$\sum_{0}^{n}{}' \cos j\beta = \frac{1}{2} \frac{\sin(n + 1/2)\beta}{\sin \beta/2}$$

yielding,

$$\sum_{0}^{n}{}' \cos^2 j\beta = \frac{1}{2}\left(n+\frac{1}{2}\right) + \frac{1}{4}\frac{\sin(2n+1)\beta}{\sin\beta}$$

one gets for $i \neq r$

$$\sum_{j=0}^{k}{}' T_j(\xi_i)\, T_j(\xi_r) = \frac{1}{4}\left[\frac{\sin\dfrac{(2k+1)(i+r-1)\pi}{2(k+1)}}{\sin\dfrac{(i+r-1)\pi}{2(k+1)}} + \frac{\sin\dfrac{(2k+1)(i-r)\pi}{2(k+1)}}{\sin\dfrac{(i-r)\pi}{2(k+1)}}\right] \tag{4.13}$$

and for $i = r$,

$$\sum_{j=0}^{k}{}' T_j^2(\xi_r) = \frac{1}{4}\left[(2k+1) + \frac{\sin\dfrac{(2k+1)(2r-1)\pi}{2(k+1)}}{\sin\dfrac{(2r-1)\pi}{2(k+1)}}\right] \tag{4.14}$$

The argument of the sine function in the numerator of (4.13) can be transformed as follows,

$$\frac{\sin\dfrac{(2k+1)(i+r-1)\pi}{2(k+1)}}{\sin\dfrac{(i+r-1)\pi}{2(k+1)}} = \frac{\sin\left((i+r-1)\pi - \dfrac{(i+r-1)\pi}{2(k+1)}\right)}{\sin\dfrac{(i+r-1)\pi}{2(k+1)}} = (-1)^{i+r}$$

and similarly for the second term one gets $(-1)^{i-r+1}$ and $(-1)^{i+r} + (-1)^{i-r+1} \equiv 0$ since $i+r$ and $i-r$ are both even or odd.

Now, for $i = r$ in (4.14), one finds that the trigonometric term is $(-1)^{2r} = 1$. Eventually, one has,

$$\sum_{j=0}^{k}{}' T_j(\xi_i)\, T_j(\xi_r) = \begin{cases} (k+1)/2 & \text{if } i = r \\ 0 & \text{if } i \neq r. \end{cases} \tag{4.15}$$

Inserting (4.15) in (4.12) the final result comes out as,

$$\boxed{\Phi(\xi_r) = f(\xi_r)} \tag{4.16}$$

meaning that the polynomial $\Phi(x)$ interpolates $f(x)$ at the Chebyshev points $\xi_1,\ \xi_2, \cdots, \xi_{k+1}$. It is important to remind that this property is true only if the Chebyshev coefficients $C_0, C_1, \cdots, C_k$ are evaluated with (4.9). □

We also see from (4.9) that this is identical to the computation of the integral (4.1) by a Gauss-Chebyshev quadrature of degree $k+1$, enhancing the close relationship between the interpolation and the Fourier expansion. A final gain with the interpolation is the fact that we can evaluate the error from the general theory of Lagrange interpolation, an alternative approach to the methods applied in Sect. 3.

Remark: It follows from this property an orthogonality relationship over the discrete set of Chebyshev points $\xi^{(k)} = \cos\theta^{(k)}$ with $\theta^{(k)} = \{\pi/2k, 3\pi/2k, \cdots, (2k-1)\pi/2k\}$. If we takes $f(x) = T_l(x)$, $l < k$, then the interpolation of a polynomial of degree $l < k$ at k points, leads not only to $\Phi(\xi_r) = f(\xi_r)$ but to the stronger result $\Phi(x) = f(x) \equiv T_l(x)$. Therefore one has in (4.9) $C_j = \delta_{jl}$ leading to,

$$\boxed{\sum_{i=1}^{k} T_r(\xi_i)\, T_l(\xi_i) = \begin{cases} k & \text{if } r = l = 0 \\ k/2 & \text{if } r = l \neq 0 \\ 0 & \text{if } r \neq l. \end{cases}} \tag{4.17}$$

Property 8. *If the coefficients $C_0,\ C_1,\ \cdots,\ C_k$ of a Chebyshev expansion are computed as :*

$$C_j = \frac{2}{k} \sum_{i=0}^{k}{}'' f(\zeta_i)\, T_j(\zeta_i), \quad j = 0,\ 1,\ \cdots,\ k \tag{4.18}$$

where $T_k(\zeta_i) = \pm 1$, that is to say the ζ_i $(i = 0, \cdots, k)$ are the the locations of the extremums of $T_k(x)$ given in (2.10) *, including the two end-points, the resulting polynomial $\Phi(x) = \sum''_{j=0}{}^{k} C_j\, T_j(x)$ of degree k interpolates $f(x)$ at the $k+1$ ζ_i, including $x = -1$ and $x = 1$.*

The summation $\sum''$ means that both the first and last terms are halved.

Remark: this means that the integrals (4.1) up to $j = k$, transformed into Fourier integrals,

$$\frac{2}{\pi}\int_0^{\pi} f(\cos\theta)\cos\, j\theta\, d\theta,\quad j = 0,\, 1,\, \cdots \tag{4.19}$$

are numerically computed with a trapezoidal rules of order k, with a regular sampling on θ between 0 and π. This is known to be very accurate for periodic functions, a property easily established with the Euler-MacLaurin formula giving a direct link between integrals and discrete sums (K. Atkinson (1989)).

Proof. Let

$$\Phi(x) = \sum_{j=0}^{j=k}{}'' C_j\, T_j(x) = \frac{1}{2}C_0\, T_0(x) + C_1\, T_1(x) + \cdots + C_{k-1}\, T_{k-1}(x) + \frac{1}{2}C_k\, T_k(x) \tag{4.20}$$

be a Chebyshev representation of $f(x),\ x \in [-1,\, 1]$, with

$$C_j = \frac{2}{k}\sum_{i=0}^{k}{}'' f(\zeta_i)\, T_j(\zeta_i),\quad j = 0,\, 1,\, \cdots,\, k \tag{4.21}$$

where $\zeta_i = \cos\frac{i\pi}{k}$. Then,

$$\Phi(\zeta_r) = \sum_{j=0}^{j=k}{}'' C_j\, T_j(\zeta_r),\quad r = 0, \cdots, k \tag{4.22}$$

and by inserting (4.21) in (4.22) we have,

$$\Phi(\xi_r) = \frac{2}{k}\sum_{j=0}^{j=k}{}''\sum_{i=0}^{k}{}'' f(\zeta_i)\, T_j(\zeta_i)\, T_j(\zeta_r),,\quad r = 0,\, \cdots,\, k$$

Using the definition of $T_k(x)$ and the values of the extremums ζ_r of $T_k(x)$ the right hand side becomes,

$$\frac{2}{k}\sum_{i=0}^{k}{}'' f(\zeta_i)\sum_{j=0}^{k}{}'' \cos\left(j\frac{i\pi}{k}\right)\, \cos\left(j\frac{r\pi}{k}\right) \tag{4.23}$$

Transforming now the product of cosines into a sum and using the results,

$$\sum_0^n{}'' \cos j\beta = \sum_0^n{}' \cos j\beta - \frac{1}{2}\cos n\beta = \frac{1}{2}\frac{\sin n\beta}{\tan \beta/2}$$

and then,

$$\sum_0^n{}'' \cos^2 j\beta = \sum_0^n{}' \cos^2 j\beta - \frac{1}{2}\cos^2 n\beta = \frac{n}{2} + \frac{1}{4}\frac{\sin 2n\beta}{\tan \beta}$$

one gets for $i \neq r$

$$\sum_{j=0}^k{}'' T_j(\zeta_i)\, T_j(\zeta_r) = \frac{1}{4}\left[\frac{\sin(i+r)\pi}{\tan\frac{(i+r)\pi}{2k}} + \frac{\sin(i-r)\pi}{\tan\frac{(i-r)\pi}{2k}}\right] \equiv 0 \tag{4.24}$$

and for $i = r$,

$$\sum_{j=0}^k{}'' T_j^2(\xi_r) = \frac{1}{4}\left[2k + \frac{\sin 2r\pi}{\tan\frac{r\pi}{k}}\right] \equiv \frac{k}{2} + \frac{\delta_{r,0}}{2} + \frac{\delta r, k}{2} \tag{4.25}$$

Eventually, one has,

$$\sum_{j=0}^k{}'' T_j(\zeta_i)\, T_j(\zeta_r) = \begin{cases} k & \text{if } i = r = 0 \text{ or } i = r = k \\ k/2 & \text{if } i = r \neq 0 \text{ or } k \\ 0 & \text{if } i \neq r. \end{cases} \tag{4.26}$$

Finally inserting (4.26) in (4.23) the final result comes out as,

$$\boxed{\Phi(\zeta_r) = f(\zeta_r)} \tag{4.27}$$

meaning that the polynomial $\Phi(x)$ interpolates $f(x)$ at the extremums points $\zeta_0, \zeta_1, \cdots, \zeta_k$, where $\zeta_0 = -1$ and $\zeta_k = 1$ are the end-points of the interpolation interval. Again, this property applied only if the coefficients up to degree k are evaluated with (4.18) and if the first and last coefficients are halved in the Chebyshev expansion. □

This interpolation is totally equivalent to computing the Chebyshev coefficients up to C_k with a composite trapezium quadrature method of degree k (same degree k, for all the coefficients $C_0, C_1, \cdots, C_k$.) Therefore it can be seen from two different perspectives: that of a numerical scheme to obtain approximate values of the c_j, $0 \leq j \leq k$ with a convenient numerical procedure, or that of an interpolation on a well chosen set of nodes expressed directly in the form of an expansion on the Chebyshev polynomials. Having included the two end-points as nodes, the interpolation error is zero at the boundaries of the interval, allowing to construct a continuous piecewise interpolation with adjacent sub-intervals.

Remark: In a similar way as for the Chebyshev point, it follows from this property an orthogonality relationship over the discrete set $\zeta^{(k)} = \cos\theta^{(k)}$ with $\theta^{(k)} = \{-1, \pi/k, 2\pi/k, \cdots, (k-1)\pi/k, 1\}$. If we takes $f(x) = T_l(x)$, $l \leq k$, then the interpolation of a polynomial of degree $l \leq k$ at $k+1$ point, leads not only to $\Phi(\zeta_r) = f(\zeta_r)$ but to the stronger result $\Phi(x) = f(x) \equiv T_l(x)$. Therefore one has in (4.18) $C_j = \delta_{jl}$ leading to,

$$\boxed{\sum_{i=0}^{k}{}'' T_r(\zeta_i)\, T_l(\zeta_i) = \begin{cases} k & \text{if } r = l = 0 \\ k/2 & \text{if } r = l \neq 0 \\ 0 & \text{if } r \neq l. \end{cases}} \tag{4.28}$$

5 Error in the numerical quadrature

In the vast majority of the situations when one wishes to approximate a function with a truncated Chebyshev expansion, the coefficients of the expansion must be determined with a numerical method. Numerical quadratures like the trapezium or the Gauss-Chebyshev have general formula to express the remainder, which depend on the value of a certain derivative at a point, in general unknown, insisde the integration interval. These expressions are very general, and in this section I derive direct formulas of the errors as a function of the truncation degree for both methods.

5.1 Gauss-Chebyshev quadrature

5.1.1 Analytical expression for the coefficients

Let the Chebyshev coefficients,

$$c_j = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)\,T_j(x)dx}{\sqrt{1-x^2}} \tag{5.1}$$

be computed with a Gauss-Chebyshev quadrature of order n as,

$$C_j = \frac{2}{n}\sum_{l=1}^{n} f(\xi_l)\,T_j(\xi_l),\quad j=0,\,1,\,\cdots,N \tag{5.2}$$

where ξ_l stands in fact for $\xi_l^{(n)}$, the i^{th} zero of $T_n(x)$. One expects to have $C_j \approx c_j$ at least for $j \leq N$, where N is of the order of n. If N is fixed, meaning the truncation level of the Chebyshev expansion is determined by a rule independent of the specific integration (5.2), it is legitimate to ask how to choose n to ensure a good accuracy in the evaluation of c_j. One feels that if $f(x)$ has many oscillations with significant amplitudes over the integration interval, a large order n will be needed. This is actually the case and in this section I will make this reasonable guess more precise.

Using the completeness of the Chebyshev basis (Davis (1975)) on the space of square integrable and continuous functions, one can write

$$f(x) = {\sum_{k=0}^{\infty}}' c_k\,T_k(x) \tag{5.3}$$

and by substituting (5.3) into (5.2) one can express the C_j as a function of the c_j. One would like to find $C_j \simeq c_j$. Combining (5.2) and (5.3) one gets,

$$C_j = \frac{2}{n} \sum_{k=0}^{\infty}{}' c_k \sum_{l=1}^{n} T_k(\xi_l)\, T_j(\xi_l), \quad j = 0,\, 1,\, \cdots, \tag{5.4}$$

With $T_k(\cos\theta) = \cos k\theta$, one can rewrite (5.4) as,

$$C_j = \frac{2}{n} \sum_{k=0}^{\infty}{}' c_k \sum_{l=1}^{n} \frac{1}{2} \left[T_{k+j}(\xi_l) + T_{|k-j|}(\xi_l)\right], \quad j = 0,\, 1,\, \cdots, \tag{5.5}$$

and with

$$\xi_l = \cos\frac{(2l-1)\pi}{2n} \tag{5.6}$$

we are left with summations of the following type,

$$\begin{aligned} S_r &= \sum_{l=1}^{l=n} \cos r\frac{(2l-1)\pi}{2n} \\ &= \sum_{l=1}^{l=n} \cos(2l-1)\alpha \\ &= \mathcal{R}\left(\sum_{l=1}^{l=n} e^{(2l-1)i\alpha}\right) \end{aligned} \tag{5.7}$$

with $\alpha = r\pi/2n$. After a little and elementary algebra one gets,

$$\begin{aligned} S_r &= \frac{\sin 2n\alpha}{2\,\sin\alpha} \\ &= \frac{\sin r\pi}{2\,\sin\dfrac{r\pi}{2n}} \end{aligned} \tag{5.8}$$

or

$$S_r = \begin{cases} n & \text{if } r = 0 \\ 0 & \text{if } 1 \le r \le 2n-1 \\ -n & \text{if } r = 2n \\ 0 & \text{if } 2n+1 \le r \le 4n-1 \\ n & \text{if } r = 4n \end{cases} \tag{5.9}$$

and the rule goes on for larger r, with S_r being non zero only for $r = 2p$ in which case $S_r = (-1)^p n$.

Remark: The use of imaginary exponential and geometric series is the systematic way of handling these trigonometric sums. For several of these sums a simpler trick can be used. Here consider the sum

$$2\sin\alpha\,(\cos\alpha + \cos 3\alpha + \cdots + \cos(2n-1)\alpha)$$

Inserting $2\sin\alpha\,\cos l\alpha = \sin[(l+1)\alpha] - \sin[(l-1)\alpha]$, there is a systematic cancellation with the negative term, and it remains only at the end the term $l = 2n-1$ and then the sum is equal to $\sin 2n\alpha$, the same as in (5.8).

Using (5.9) in (5.5), successively with $r = k+j$ and $r = |k-j|$, one finds the important result,

$$\boxed{C_j = c_j - c_{2n-j} - c_{2n+j} + c_{4n-j} + c_{4n+j} + (-1)^p c_{2np-j} + (-1)^p c_{2np+j} + \cdots} \tag{5.10}$$

which relates the approximate value C_j of a numerically evaluated Chebyshev coefficient with a Gauss-Chebyshev method of order n to its true value c_j. Essentially (5.10) tells us that C_j is polluted by projection of higher order terms, in a way very similar to aliasing from frequencies above the Nyquist limit in harmonic analysis. Eq. 5.10 is very general and can be specialised by assuming that the c_j decrease with j, which is a very common situation. Then if $j \le n$ (order of integration larger than, or equal to, the highest degree of the expansion), one has in (5.10), $C_j \approx c_j - c_{j+2(n-j)}$, as the other terms are much smaller. Explicitly,

$$C_{n-1} \approx c_{n-1} - c_{n+1} \tag{5.11a}$$
$$C_{n-2} \approx c_{n-2} - c_{n+2} \tag{5.11b}$$
$$\vdots \tag{5.11c}$$
$$C_{n-k} \approx c_{n-k} - c_{n+k} \tag{5.11d}$$

Therefore, the accuracy of the coefficients with $j < n$ depends on how fast the c_j decreases with increasing j.

Also, with the same order of integration,

$$C_n \approx 0 \tag{5.12a}$$
$$C_{n+1} \approx c_{n+1} - c_{n-1} \tag{5.12b}$$
$$C_{n+2} \approx c_{n-2} - c_{n+2} \tag{5.12c}$$
$$\vdots \tag{5.12d}$$

showing that as soon as $j \geq n$, the numerically computed coefficients are totally wrong.

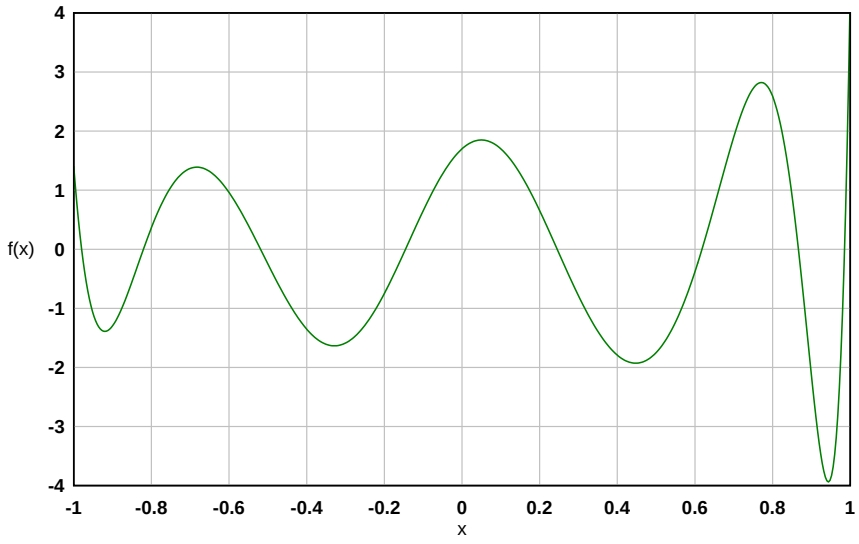


FIGURE 5.1: Test function (5.13) used to illustrate the properties of numerical quadrature applied to the evaluation of the Chebyshev coefficients.

The following numerical example built on a purposely test function, (Fig. 5.1)

$$f(x) = 2T_8(x) + T_9(x) + 0.6T_{10}(x) + 0.4T_{11}(x) + 0.3T_{12}(x) + 0.1T_{13}(x) \tag{5.13}$$

confirms fully the expected behaviour of the numerical quadrature. The order of the integration has been taken as $n = 10$ for a Gauss-Chebyshev quadrature. Results are given in Table 5.1, with the true values of the coefficients, the result of the Gauss-Chebyshev numerical quadrature of order 10 and then of order $n = 20$ in the last column. One can check that the results match exactly (5.11) and (5.12).

TABLE 5.1: Computation of the Chebyshev coefficients with a Gauss-Chebyshev quadrature. The expansion goes to degree 15 and the order of integration is $n = 10$. The columns are the rank of the coefficients, the true value for the test function, the numerical values based on a Gauss-Chebyshev quadrature of order $n = 10$ and then of order $n = 20$. One sees for example that $C_{10} = 0.0$ instead of 0.6, in agreement with (5.12).

j	c_j	C_j	C_j^*
0	0.0	-0.19929E-14	0.39968E-15
1	0.0	-0.20983E-14	0.42188E-15
2	0.0	-0.41855E-14	0.10436E-14
3	0.0	-0.25202E-14	0.66613E-15
4	0.0	-0.46851E-14	0.93259E-15
5	0.0	-0.28089E-14	0.22204E-15
6	0.0	-0.44631E-14	0.44409E-16
7	0.0	-0.10000E+00	-0.39968E-15
8	2.0	0.17000E+01	0.20000E+01
9	1.0	0.60000E+00	0.10000E+01
10	0.6	-0.15846E-14	0.60000E+00
11	0.4	-0.60000E+00	0.40000E+00
12	0.3	-0.17000E+01	0.30000E+00
13	0.1	0.10000E+00	0.10000E+00
14	0.0	0.38525E-14	-0.10880E-14
15	0.0	0.24536E-14	-0.49405E-15

5.1.2 Error in the truncated expansion

Let $f(x)$ be a function expandable with uniform convergence over the Chebyshev polynomials as,

$$f(x) = \sum_{k=0}^{\infty}{}' \, c_k \, T_k(x) \tag{5.14}$$

and

$$f_n(x) = \sum_{k=0}^{n}{}' \, c_k \, T_k(x) \tag{5.15}$$

the corresponding truncated expansion to degree n. I also assume that the magnitude of the c_k decreases quickly enough with k so that the truncation error $(f_n(x) - f(x))$ is primarily set by the first neglected term $c_{n+1}T_{n+1}(x)$.

Let also the c_k be approximated by C_k computed with the above Gauss-Chebyshev quadrature of order $n+1$, so that the expansion,

$$\phi_n(x) = {\sum_{k=0}^{n}}' \, C_k \, T_k(x) \tag{5.16}$$

is equivalent to a polynomial interpolation of the $f(x)$ at the Chebyshev points. Using now (5.12) and (5.11) it is easy to evaluate the actual error in the truncated expansion (5.32) based on the approximate coefficients C_k. Since the c_k decreases quickly with k one can focus on what happens in the vicinity of $k = n$. With the quadrature at the order $n+1$ one has,

$$C_{n-1} \approx c_{n-1} - c_{n+3} \tag{5.17a}$$
$$C_n \approx c_n - c_{n+2} \tag{5.17b}$$
$$C_{n+1} \approx c_{n+1} - c_{n+1} \equiv 0 \tag{5.17c}$$

yielding for the main terms contributing to the error,

$$\phi_n(x) - f(x) \approx -c_{n+2}T_n - c_{n+1}T_{n+1} - c_{n+2}T_{n+2} \tag{5.18a}$$
$$= -c_{n+1}T_{n+1} - 2x\, c_{n+2}T_{n+1} \tag{5.18b}$$

instead of $-c_{n+1}T_{n+1} - c_{n+2}T_{n+2}$ for $f_n(x) - f(x)$. With $|c_{n+1}| \gg |c_{n+2}|$, the error remains of the order of $|c_{n+1}|$ and in most cases of interest the interpolation at the Chebyshev points expanded on the Chebyshev polynomials is very similar to a genuine Chebyshev expansion. Naturally the interpolating polynomial can be obtained by the standard Lagrange interpolation technique or with the divided difference method. However, the best way to have it expressed directly with the Chebyshev polynomials up to the degree n is then to compute the coefficients with a Gauss-Chebyshev quadrature of order $(n+1)$.

5.2 Composite trapezium quadrature

When the integral giving the Chebyshev coefficients is regularised with $x = \cos\theta$ one has,

$$c_j = \frac{2}{\pi}\int_0^{\pi} f(\cos\theta)\,\cos j\theta d\theta \tag{5.19}$$

which can be evaluated with a composite trapezium method of order n (n intervals or $n+1$ nodes, including the end-points) as,

$$C_j = \frac{2}{n}\sum_{l=0}^{n}{}'' f(\zeta_l)\,T_j(\zeta_l),\quad j = 0,\,1,\,\cdots,N \tag{5.20}$$

where $\zeta_l = \cos\theta_l$ stands in fact for $\zeta_l^{(n)}$, the i^{th} extremum of $T_n(x)$, with,

$$\theta_l = \frac{l\pi}{n}\quad l = 0,1,\cdots,n \tag{5.21}$$

As before, considering that $f(x)$ can be expanded on the Chebyshev polynomials as (5.3), which leads to

$$C_j = \frac{2}{n}\sum_{k=0}^{\infty}{}' c_k \sum_{l=0}^{n}{}'' T_k(\zeta_l)\,T_j(\zeta_l),\quad j = 0,\,1,\,\cdots,N \tag{5.22}$$

With $T_k(\cos\theta) = \cos k\theta$, one can rewrite (5.22) as,

$$C_j = \frac{2}{n}\sum_{k=0}^{\infty}{}' c_k \sum_{l=0}^{n}{}'' \frac{1}{2}\left[\cos(k+j)\zeta_l + \cos(|k-j|)\zeta_l\right],\quad j = 0,\,1,\,\cdots, \tag{5.23}$$

and with

$$\zeta_l = \cos\frac{l\pi}{n} \tag{5.24}$$

we have just to evaluate expressions similar to,

$$S_r = \sum_{l=0}^{l=n}{}'' \cos l\alpha,\quad \alpha = \frac{r\pi}{n},\quad r \text{ integer} \tag{5.25}$$

or,

$$S_r = \mathcal{R}\left(\frac{1}{2} + e^{i\alpha} + e^{2i\alpha} + \cdots + e^{i(n-1)\alpha} + \frac{1}{2}e^{i\,n\alpha}\right) \tag{5.26}$$

$$\begin{aligned} S_r &= \frac{\cos\dfrac{\alpha}{2}\,\sin n\alpha}{2\,\sin\dfrac{\alpha}{2}} \\ &= \frac{\cos\dfrac{r\pi}{2n}\,\sin r\pi}{2\,\sin\dfrac{r\pi}{2n}} \end{aligned} \tag{5.27}$$

yielding,

$$S_r = \begin{cases} n & \text{if } r = 0 \\ 0 & \text{if } 1 \leq r \leq 2n-1 \\ n & \text{if } r = 2n \\ 0 & \text{if } 2n+1 \leq r \leq 4n-1 \\ n & \text{if } r = 4n \end{cases} \tag{5.28}$$

and the rule goes on for larger r, with S_r being non zero only for $r = 2p$ in which case $S_r = n$.

Using (5.28) in (5.23), successively with $r = k + j$ and $r = |k - j|$, one finds the important result,

$$\boxed{C_j = c_j + c_{2n-j} + c_{2n+j} + c_{4n-j} + c_{4n+j} + \cdots + c_{2np-j} + c_{2np+j} + \cdots} \tag{5.29}$$

The similarity between (5.10) and (5.29) is striking.

$$C_{n-1} \approx c_{n-1} + c_{n+1} \tag{5.30a}$$
$$C_{n-2} \approx c_{n-2} + c_{n+2} \tag{5.30b}$$
$$\vdots \tag{5.30c}$$
$$C_{n-k} \approx c_{n-k} + c_{n+k} \tag{5.30d}$$

TABLE 5.2: Computation of the Chebyshev coefficients with a composite trapezium quadrature. The expansion goes to degree 15 and the order of integration is $n = 10$. The columns are the rank of the coefficients, the true values for the test function, the numerical values based on a numerical quadrature with composite trapezium method of order $n = 10$ and then $n = 20$. One sees for example that $C_{10} = 1.2$ instead of 0.6, in agreement with (5.31).

j	c_j	C_j	C_j^*
0	0.0	0.99920E-16	-0.77161E-15
1	0.0	0.28866E-15	-0.87708E-15
2	0.0	0.99920E-15	-0.18097E-14
3	0.0	0.37748E-15	-0.92149E-15
4	0.0	0.64393E-15	-0.18097E-14
5	0.0	0.19984E-15	-0.15876E-14
6	0.0	-0.42188E-15	-0.25646E-14
7	0.0	0.10000E+00	-0.61062E-15
8	2.0	0.23000E+01	0.20000E+01
9	1.0	0.14000E+01	0.10000E+01
10	0.6	0.12000E+01	0.60000E+00
11	0.4	0.14000E+01	0.40000E+00
12	0.3	0.23000E+01	0.30000E+00
13	0.1	0.10000E+00	0.10000E+00
14	0.0	-0.30864E-14	0.58842E-15
15	0.0	-0.13989E-14	0.72164E-15

Therefore, the accuracy of the coefficients with $j < n$ depends on how fast the c_j decreases with increasing j.

Also, with the same order of integration,

$$C_n \approx 2c_n \tag{5.31a}$$
$$C_{n+1} \approx c_{n+1} + c_{n-1} \tag{5.31b}$$
$$C_{n+2} \approx c_{n-2} + c_{n+2} \tag{5.31c}$$
$$\vdots \tag{5.31d}$$

showing that as soon as $j \geq n$, the numerically computed coefficients are totally wrong. Incidentally, with $C_n = 2c_n$, one understands why the expansion based on a Trapezium composite quadrature is better when the last coefficient is halved.

Using also (5.13) for the trial function, numerical results are given in Table 5.2, with the true values of the coefficients, the result of the numerical quadrature of order 10 and then of order $n = 20$ in the last column. One can check that the results match exactly (5.30) and (5.31).

5.2.1 Error in the truncated expansion

Using exactly the same notations and principles as in Sect. 5.1.2 but with,

$$\phi_n(x) = \sum_{k=0}^{n}{}'' \, C_k \, T_k(x) \tag{5.32}$$

(the C_n coefficient in the sum is also halved) and provided one computes the C_k with the trapezium quadrature of order n (and not $n+1$ as with the Gauss-Chebyshev), one has

$$C_{n-1} \approx c_{n-1} + c_{n+1} \tag{5.33a}$$
$$C_n \approx 2c_n \tag{5.33b}$$
$$C_{n+1} \approx c_{n+1} + c_{n-1} \tag{5.33c}$$

Therefore the error produced when using the interpolation at the Chebyshev-Lobatto points instead of a full Chebyshev expansion is given by,

$$\phi_n(x) - f(x) \approx c_{n+1}T_{n-1} - c_{n+1}T_{n+1} - c_{n+2}T_{n+2} \tag{5.34a}$$
$$= -c_{n+1}(T_{n+1} - T_{n-1}) \tag{5.34b}$$
$$= -c_{n+1}2(1-x^2)U_{n-1} \tag{5.34c}$$

where $U_n(x)$ is the Chebyshev polynomial of second kind of degree n. The polynomials $2(1-x^2)U_n(x)$ of degree $n+2$ are plotted in Fig. 5.2. They form a set of polynomials whose zeros are the Chebyshev-Lobatto points, including the end-points ± 1. One sees in the plot that their magnitude is always ≤ 1, and then the error in the Chebyshev-Lobatto interpolation is of the order of $2c_{n+1}$, or at most twice the error that would be obtained with a standard Chebyshev expansion to the same degree. This is the price to pay to ensure the continuity, but at the same time one retains most of the nice properties of the Chebyshev expansion, with an error relatively evenly distributed and nothing equivalent of the Runge effect at large degree. Though not formally established here, it is clear that we will have a uniform convergence with increasing degrees.

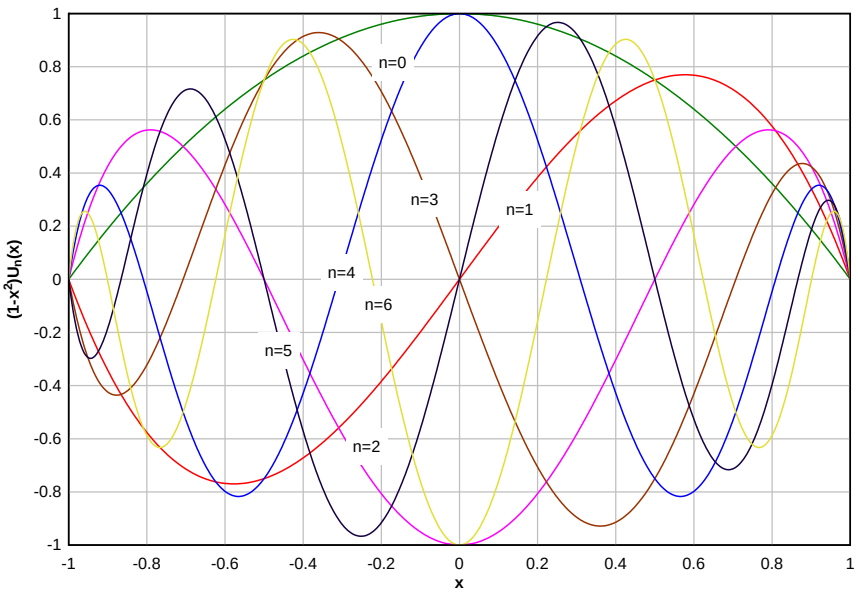


FIGURE 5.2: Plot of the function $(1-x^2)U_n(x)$, where $U_n(x)$ is the Chebyshev polynomial of second kind. The remainder of a Gauss-Lobatto interpolation is closely related to this function.

5.3 A worst case function

As long as the coefficients c_k of the Chebyshev expansion decrease reasonably fast with k, the different methods discussed above work nicely and one has a good control of the error. Consider as a final illustration of this section a sort of worst case function, where the coefficients c_k are constant. This is akin to the white noise signal is Fourier analysis, where all frequencies make up the signal with equal amplitudes. This is a very good function to experiment with to evidence the differences between the full Chebyshev approximation and the two kinds of interpolation, according to the truncation level and the orders of integration used

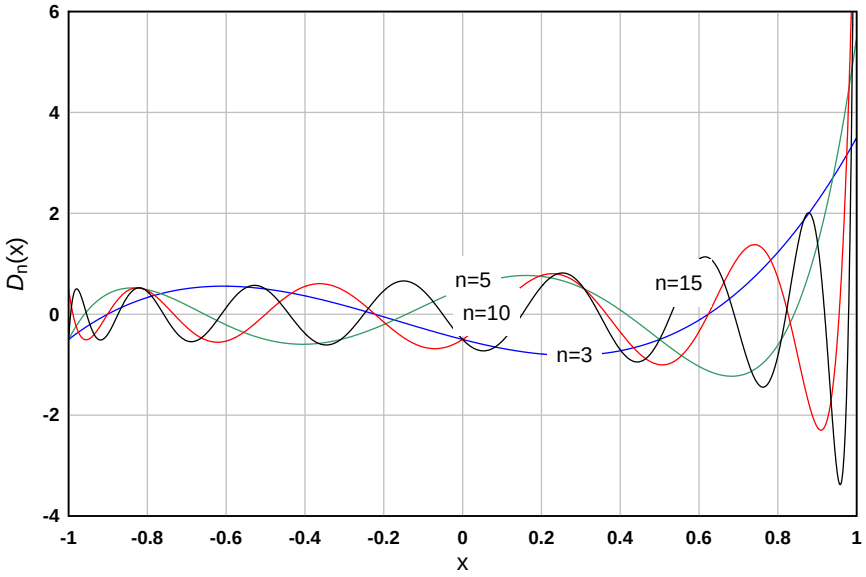


FIGURE 5.3: Sample plots of the Dirichlet kernel $\mathcal{D}_n(x)$ for n = 3, 5, 10 and 15. This function has a Chebyshev expansion with coefficients $c_j = 1$ up to $j = n$ and zero for $j > n$.

If we take $c_0 = 0.5$ and $c_k = 1$ for $1 \leq k \leq n$ up to a certain degree $n = n_{\max}$, we have a function like,

$$\mathcal{D}_n(x) = {\sum_{k=0}^{k=n}}' T_k(x) \tag{5.35}$$

and with $x = \cos\alpha$, this is equivalent to,

$$\mathcal{D}_n(\alpha(x))) = \frac{1}{2}\frac{\sin\left(n+1/2\right)\alpha}{\sin\frac{\alpha}{2}} \tag{5.36}$$

which is known in Fourier analysis under the name of the Dirichlet kernel. It can be explicitly computed as a function of x with $\alpha = \arccos(x)$. Numerically a little care is needed to handle

the singularity at $x = 1$ where one has $\mathcal{D}_n(1) = n + 1/2$. Sample plots of the Dirichlet kernel as shown in Fig. 5.3.

I have computed the coefficients C_j up to $j = 10$ of the Chebyshev expansion and that of the Chebyshev representation of a Gauss-Chebyshev and a Gauss-Lobatto interpolation of $\mathcal{D}_{15}(x)$, with $d = 10, 12, 14$ for the degrees of the interpolation. The corresponding coefficients are given in col. $2 - 4$ of Tables 5.3-5.5. Knowing that $c_j = 1$ for $j \leq 15$, it is easy to see that the values are in full agreement with the general equations (5.10) and (5.29). With this worst case example, the truncation at $n = 10$ for the coefficients is the main source of error, since there is power in the signal at higher degree. Therefore the errors between the function and its representations are very similar for the three expansions, as illustrated in Fig. 5.4 for $d = 12$.

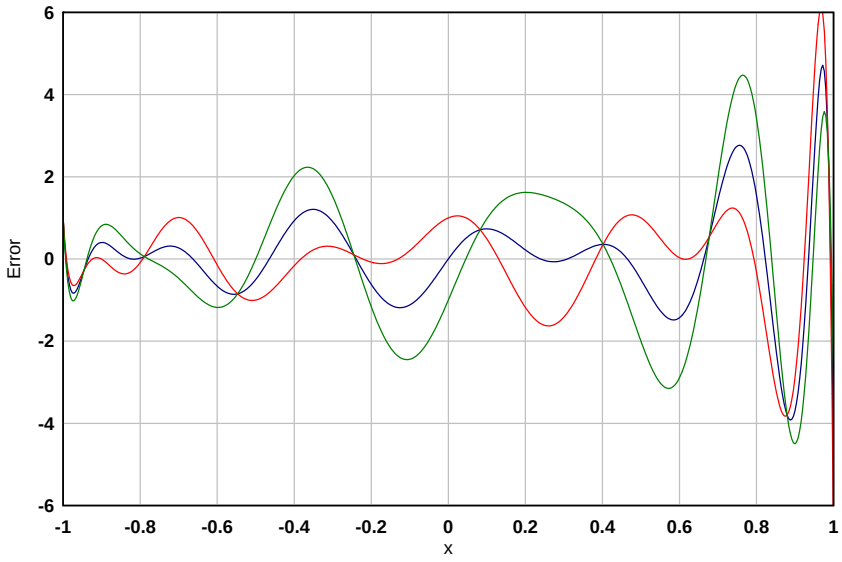


FIGURE 5.4: Error in the truncated ($n = 10$) Chebyshev approximation and interpolations of the Dirichlet kernel $\mathcal{D}_{15}(x)$. Blue curve: standard Chebyshev approximation, red curve: Gauss-Chebyshev interpolation, green curve: Gauss-Lobatto interpolation.

TABLE 5.3: Computation of the Chebyshev coefficients for the Dirichlet kernel $\mathcal{D}_{15}(x)$. The expansion is computed to degree $n = 10$ with an exact quadrature (2nd col.), then for a Gauss-Chebyshev and Gauss-Lobatto interpolation of degree 10. The values of the coefficients are in perfect agreement with (5.10) and (5.29).

j	c_j	C_j	C_j^*
0.	0.50000E+00	0.50000E+00	0.50000E+00
1.	0.10000E+01	0.10000E+01	0.10000E+01
2.	0.10000E+01	0.10000E+01	0.10000E+01
3.	0.10000E+01	0.10000E+01	0.10000E+01
4.	0.10000E+01	0.10000E+01	0.10000E+01
5.	0.10000E+01	0.46518E-14	0.20000E+01
6.	0.10000E+01	0.42022E-14	0.20000E+01
7.	0.10000E+01	0.34750E-14	0.20000E+01
8.	0.10000E+01	0.25396E-14	0.20000E+01
9.	0.10000E+01	0.26548E-14	0.20000E+01
10.	0.10000E+01	0.11337E-14	0.20000E+01

TABLE 5.4: Computation of the Chebyshev coefficients for the Dirichlet kernel $\mathcal{D}_{15}(x)$. The expansion is computed to degree $n = 10$ with an exact quadrature (2nd col.), then for a Gauss-Chebyshev and Gauss-Lobatto interpolation of degree 12. The values of the coefficients are in perfect agreement with (5.10) and (5.29).

j	c_j	C_j	C_j^*
0.	0.50000E+00	0.50000E+00	0.50000E+00
1.	0.10000E+01	0.10000E+01	0.10000E+01
2.	0.10000E+01	0.10000E+01	0.10000E+01
3.	0.10000E+01	0.10000E+01	0.10000E+01
4.	0.10000E+01	0.10000E+01	0.10000E+01
5.	0.10000E+01	0.10000E+01	0.10000E+01
6.	0.10000E+01	0.10000E+01	0.10000E+01
7.	0.10000E+01	0.10000E+01	0.10000E+01
8.	0.10000E+01	0.10000E+01	0.10000E+01
9.	0.10000E+01	-0.58611E-14	0.20000E+01
10.	0.10000E+01	-0.42778E-14	0.20000E+01

TABLE 5.5: Computation of the Chebyshev coefficients for the Dirichlet kernel $\mathcal{D}_{15}(x)$. The expansion is computed to degree $n = 10$ with an exact quadrature (2nd col.), then for a Gauss-Chebyshev and Gauss-Lobatto interpolation of degree 14. The values of the coefficients are in perfect agreement with (5.10) and (5.29).

j	c_j	C_j	C_j^*
0.	0.50000E+00	0.50000E+00	0.50000E+00
1.	0.10000E+01	0.10000E+01	0.10000E+01
2.	0.10000E+01	0.10000E+01	0.10000E+01
3.	0.10000E+01	0.10000E+01	0.10000E+01
4.	0.10000E+01	0.10000E+01	0.10000E+01
5.	0.10000E+01	0.10000E+01	0.10000E+01
6.	0.10000E+01	0.10000E+01	0.10000E+01
7.	0.10000E+01	0.10000E+01	0.10000E+01
8.	0.10000E+01	0.10000E+01	0.10000E+01
9.	0.10000E+01	0.10000E+01	0.10000E+01
10.	0.10000E+01	0.10000E+01	0.10000E+01

6 Numerical illustrations

To illustrate some of the properties introduced in the previous sections, a sample function has been represented with different kinds of Chebyshev expansions. The function equation is given in (6.1), where t can been considered as a time in days. It is a very simplified version of a typical multiperiodic function of the kind frequently seen in solar system dynamics, with a first term having a large amplitude compared to all the others. The expansion is computed over 40 days with time intervals of 8 days, so that the sub-interval is less than the period of the main component, but larger than the other two periods of the minor components.

$$f(t) = \sin(2\pi\, t/10) + 0.2\sin(2\pi\, t/7 + 0.5) + 0.05\sin(2\pi\, t/4 - 0.5) \tag{6.1}$$

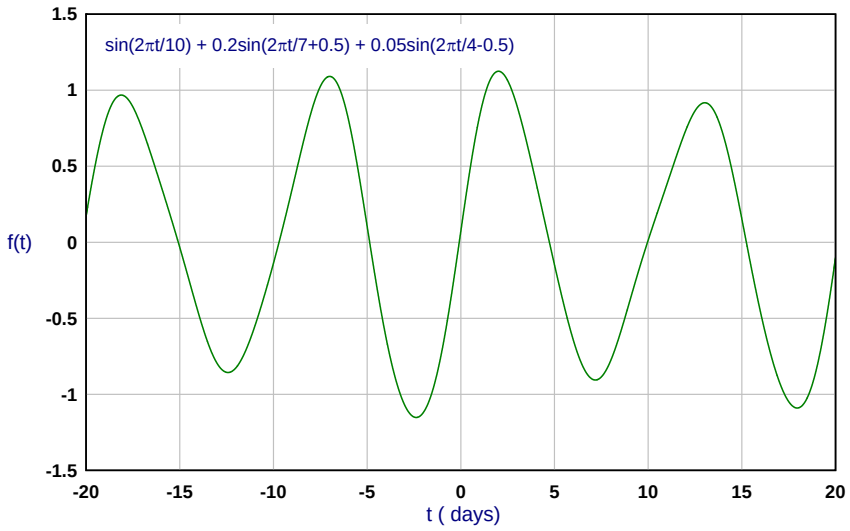


FIGURE 6.1: Test function (6.1) used to illustrate the properties of numerical quadrature applied to the evaluation of the Chebyshev coefficients.

6.1 Interpolation at Chebyshev points

The first example shown in Fig. 6.2, compares a truncated Chebyshev expansion at degree $n = 10$ to the interpolation at the Chebyshev points $\xi_k^{(11)}, 1 \le k \le 11$. Both sets of Chebyshev coefficients have been computed with a Gauss-Chebyshev quadrature, respectively of order 15 and 11. As expected from the discussion in Sect. 5.1.2, there are no significant differences between the two representations, both giving a maximum error of the order of 3×10^{-4} with a near quasi-oscillatory pattern foreseen for a near minimax approximation. More with this example, the accuracy can be predicted with the magnitude of c_{11} given by (3.20).

In this trial function, it is not immediately clear which of the three periodic components will be the dominant source of the error. The amplitudes decrease from the first to the third component, but the period of each component decreases as well, making the representation more difficult over each sub-interval of 8 days. It is easy to see that the small period takes the lead to the amplitude and that the final error comes from the third component and with $k = 11$ and $\Delta t/T = 2$ in (3.20) one has $c_k = 6.3 \times 10^{-3}$ and with the amplitude of 0.05 this gives an error for this component of $\sim 3 \times 10^{-4}$, whereas it is respectively $\sim 6 \times 10^{-7}$ and $\sim 4 \times 10^{-6}$ for the first and second periodic term. This is a very common feature in Solar system ephemerides, since the motion is in fact the superimposition of many periodic terms covering a large range of frequencies and amplitudes.

The choices of the degree and of the size of the sub-interval are not necessarily driven by the leading term in amplitude but rather by few less significant terms at higher frequency. This is normally not too difficult to find the optimum from a frequency analysis of the time series produced by a numerical integration. Since the long period terms are not a source of worry, this analysis can be performed on a rather small sample, with good time resolution to locate the short period terms.

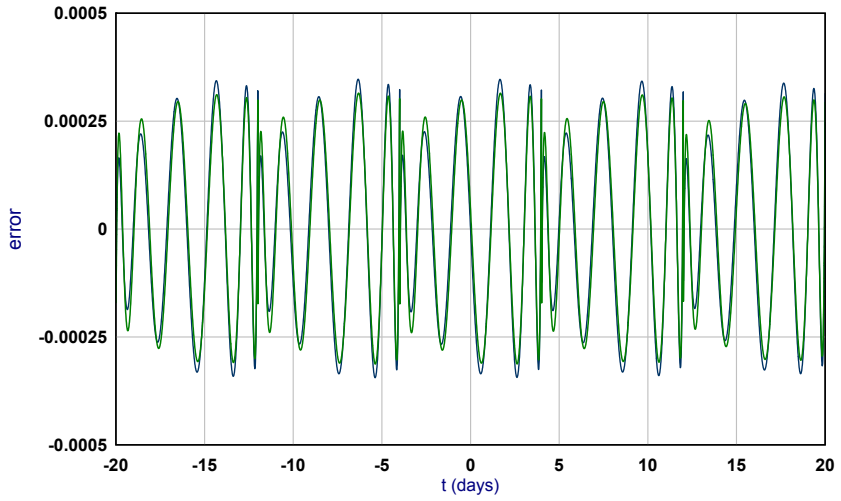


FIGURE 6.2: Chebyshev expansion truncated at degree 10 of (6.1), over 40 days with 5 interval of 8 days each. The green curve shows the expansion with accurate numerical evaluation of the theoretical coefficients, while the blue one corresponds to an interpolation at the Chebyshev points, or equivalently the use of a Gauss-Chebyshev quadrature of degree 11 to evaluate the Chebyshev coefficients.

6.2 Interpolation at Lobatto points

A similar comparison is shown in Fig. 6.3, by using a trapezium quadrature of order 10 to obtain directly the Chebyshev representation of the function interpolated at the Chebyshev-Lobatto points $\zeta_k^{(10)}, 0 \leq k \leq 10$. From (5.34) one knows that the error is typically twice as large as that with the standard Chebyshev expansion near the centre of the interval (here the centre of each sub-interval), but smaller at the outer boundaries, being exactly zero at the end-points. The representation is then continuous through the sub-intervals. These expectations are fully met in this illustrative example.

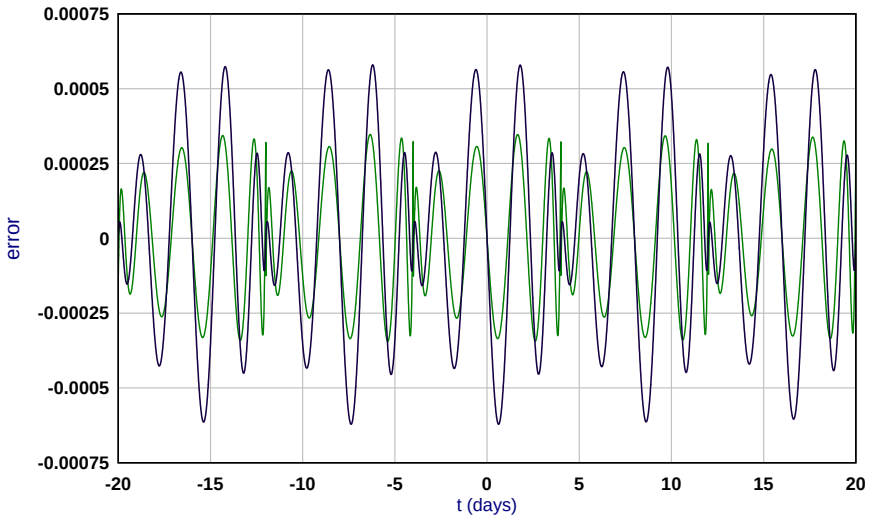


FIGURE 6.3: Comparison of two quasi Chebyshev expansions of degree 10, based on Gauss-Chebyshev (green) and Trapezium quadrature (black) or order 11.

The different behaviour of the two representations at the border of one sub-interval is plotted in Fig. 6.4 with a zoom around $t = 4$ in Fig. 6.3. The continuity with the Chebyshev-Lobatto representation is perfectly achieved. The apparent continuity of the slope is a chance effect here and is not a general property.

6.3 Least square fitting

In the following two plots the Chebyshev coefficients have been computed with a least squares fitting on sampled values of the function $f(t)$. To illustrate the importance of the weighting, or equivalently that of the sampling, I have taken two different samplings: a uniform sampling over every sub-interval, with the two end points, and a sampling at the Chebyshev points within

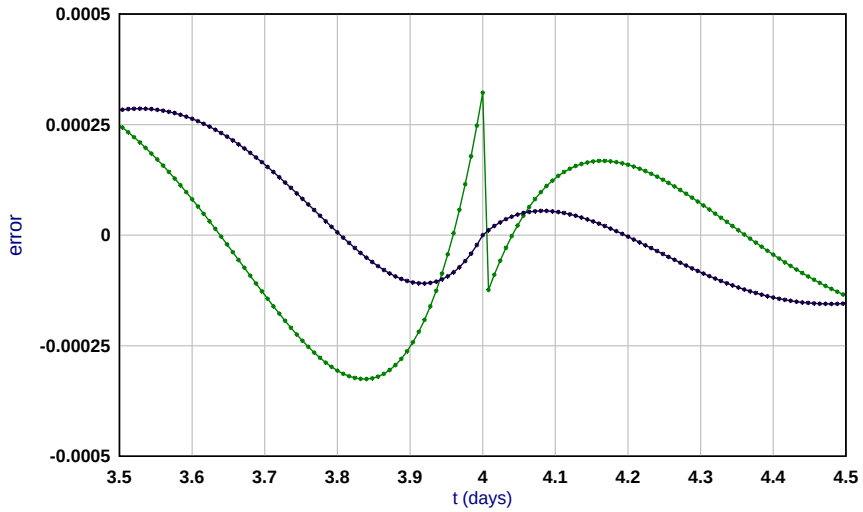


FIGURE 6.4: Zoom on Fig.6.3 at the boundary of an interval, showing the discontinuity for the Gauss-Chebyshev quadrature (same as interpolation at the Chebyshev points) and the continuity with the composite trapezium (same as interpolation at the Gauss-Lobatto points).

each sub-interval. The unweighted uniform sampling is equivalent to the standard Lagrange interpolation at regularly distributed points. As shown in Sect. 1.4.2, the choice of the Chebyshev points is a computationally stable solution to weigh the condition equations in proportion of $1/\sqrt{1-x^2}$ and set the least squares problem properly. The expansion is taken for a degree $n = 10$, meaning there are 11 unknowns in the fit and I took successively $m = 15$ and $m = 30$ nodes.

Results are shown in Fig. 6.6–6.5, respectively for $m = 15$ and $m = 30$ samples. No big surprise: a least squares fitting is an acceptable procedure to determine a Chebyshev expansion with a sampling at the Chebyshev points, and definitely not as good with a regular distribution. Using a regular distribution produces a Runge effect at each boundaries of the sub-intervals, with a significant increase of the error when the problem is slightly overdetermined. The errors are also less regularly distributed than with the Chebyshev points. Using more nodes seems to solve the problem, although no clear rule can be provided as to how many points to use. On the other hand, with Chebyshev points, there is no issue regarding the number of points, and if we specifically takes $m = n + 1$, one recovers the interpolation at the Chebyshev points discussed in the previous sections.

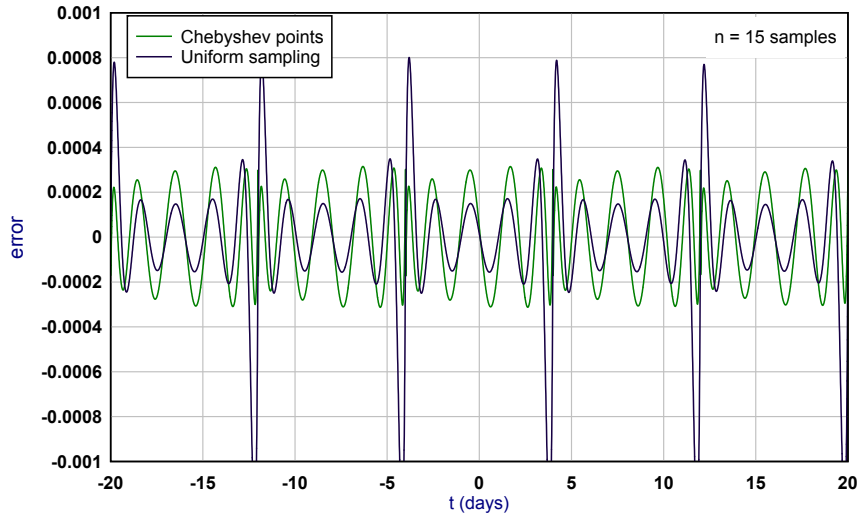


FIGURE 6.5: Error in the Chebyshev expansion of degree 10 based on Least-Squares methods using 15 data points at the Chebyshev points (green curve) or with uniform sampling over $(-1, 1)$ (blue curve).

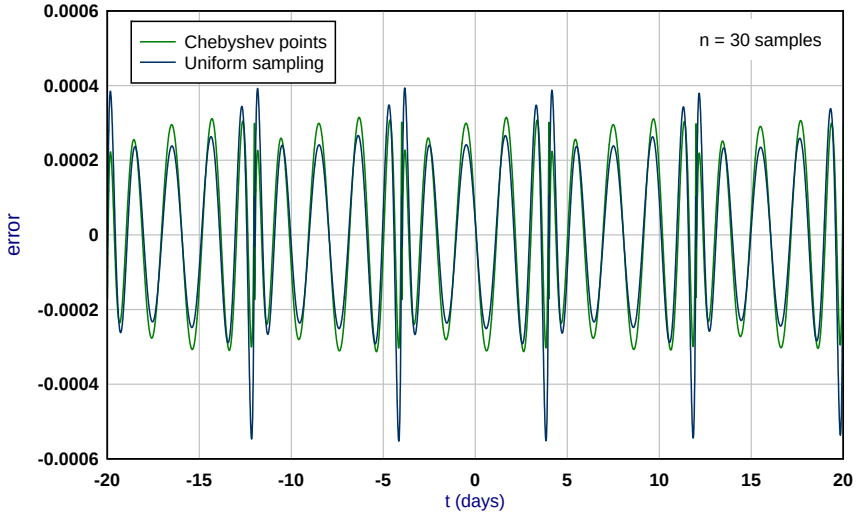


FIGURE 6.6: Error in the Chebyshev expansion of degree 10 based on Least-Squares methods using 30 data points at the Chebyshev points (green curve) or with uniform sampling over $(-1, 1)$ (blue curve).

7 Comparison to splines

Spline (specifically their cubic variety) is a very common and powerful piecewise interpolation tool producing continuous and twice continuously differentiable functions by gluing together third degree polynomials. The principle is described in any modern textbook of numerical analysis and useful information is also easily available on-line. Basically the cubic polynomials are pieced together in such a way that their values and those of the first two derivatives coincide at a set of knots. Unlike the usual Lagrange interpolation or the polynomial approximation discussed in the earlier sections of this note, there is no general results regarding the remaining error when approximating a continuous function ($f(x)$ on the interval $[a, b]$ with a cubic spline $S(x) \in \mathcal{C}^2(a, b)$ at the nodes $(a = x_0, x_1, \cdots, x_n = b)$. Since by increasing the number of nodes, one finds usually a smaller error, to perform a sensible comparison with the Chebyshev approximation either the standard or the continuous form), one needs to set a correspondence between the number of nodes and the order of the Chebyshev approximation. I consider only the cost of evaluation of the approximated functions, assuming one needs to carry out this evaluation at many points so that this cost is much higher than the evaluation of the spline or that of the Chebyshev coefficients.

Let $\mathcal{I} = [a, b]$ the interval of interest, n the number of nodes of the spline, d the degree of the Chebyshev approximation, and g the number of sub-intervals used in the coverage of $\mathcal{I}$ for the Chebyshev approximation. In addition to the nodes x_i and the values of $y_i = f(x_i)$, there will be n additional coefficients $s_1, s_2, \cdots s_n$ in the spline approximation. For the Chebyshev representation the total number of coefficients will be $N = g(d + 1)$ and the users must be also given the boundaries of each sub-intervals. A fair comparison is then to take n of the order of N. Obviously the error achieved by the approximation is not the sole criteria that must be considered. The continuity and differentiability to 2nd order of the spline is valuable asset not so easily quantifiable. I would say that the spline beats Chebyshev if the errors are comparable and as long as $N \leq 2g(d + 1)$, meaning that with similar results, I accept to pay more for a spline in terms of number of coefficients, given its smoothness over the whole interval.

7.1 Comparison within a simple interval

By *simple interval* it must be understood that the interval is not subdivided into smaller pieces to perform the Chebyshev approximation. Therefore the function $f(x)$ over the interval $[a, b]$ is expanded in Chebyshev polynomial to degree n over this interval, and similarly it is interpolated with a cubic spline, using k knots. I took,

$$f(x) = \sin(2\pi\, x + 0.6), \quad -1 \leq x \leq 1 \tag{7.1}$$

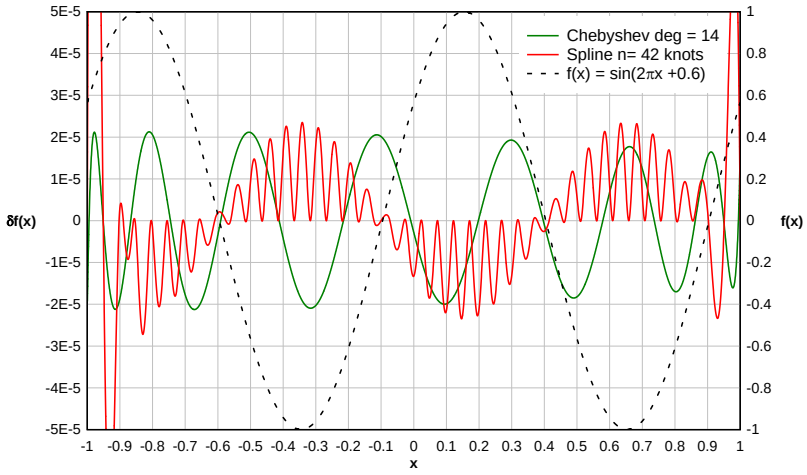


FIGURE 7.1: Comparison of a Chebyshev approximation (degree $d = 14$, green curve) to a spline interpolation (42 knots, red curve) for the function $\sin(2\pi x + 0.6)$ with x in $[-1, 1]$ (dashed curve, right scale). The error in the Chebyshev representation is in agreement with Fig 3.1 and the number of knots in the Spline has been so selected to have a similar error in the interpolation. One sees that the spline generates a large number of oscillation in the residuals (meaning that its derivative would be a very poor approximation to $f'(x)$) and that the error grows very much near the boundaries. There is also a marked relationship between the errors in the spline interpolation and the phase of the function interpolated, not a very nice feature.

The interval covers two full periods of the function and using Fig. 3.1 one sees that with $n = 14$ one expects an error of the order of 5×10^{-5} (the first neglected term must be read on the curve labelled $n = 15$ in the plot). For the spline interpolation, with trials and errors I found that one needs to take a number of knots $k \sim 42$ to obtain a similar average error, except near the boundaries, where it is much larger.

This is shown in Fig. 7.1, with the function plotted (dashed curve and right scale) and the errors respectively in the Chebyshev expansion and the spline interpolation. The error distribution over the interval is markedly different between the two methods, much in favour of the Chebyshev expansion: much smaller number of oscillations, fairly uniform distribution of the error, no growth near the boundaries, and a number of coefficients three times as small as with the spline interpolation. Over a single interval, both representations are continuous and differentiable at least twice. However, the strong oscillatory feature of the spline, makes it totally unsuitable to be a good approximation of the first derivative of the function $f(x)$. Differentiating the Chebyshev itself would not be the recommended solution, but at least this remains an acceptable solution.

7.2 Comparison within a composite interval

The potential advantage of splines, namely their smoothness and differentiability, appears more neatly when one considers large intervals, upon which the Chebyshev approximation must broken down into a set of smaller sub-intervals, each one being independent of the others. I have already discussed how to ensure at least a continuity at the boundaries and showed the cost in terms of accuracy. For the sake of comparison, the approximation is performed on the same test function as in Sect. 6 with Eq.6.1, over the same interval of 40 days divided into 5 sub-intervals of 8 days each. This is just an illustration not sufficient to draw fully general conclusions, but showing at least features that should be shared by approximations on typical quasi-periodic functions. All the following examples compare the precision achieved with a Chebyshev approximation to that of a spline interpolation, with the degree and number of nodes being adjusted to end up with similar mean error over most of the interval.

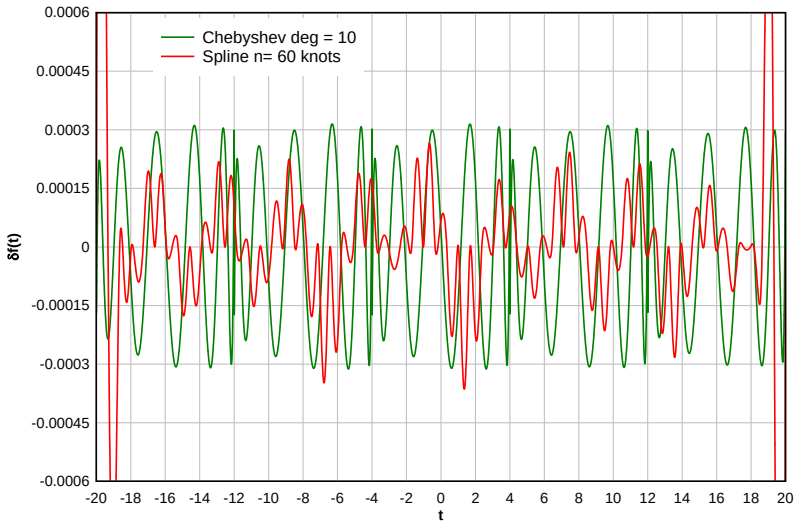


FIGURE 7.2: Comparison of the error of a Chebyshev approximation (degree $d = 10$, green curve) to a spline interpolation (60 knots, red curve) for the function defined by (6.1) over a period of 40 days and using 5 subintervals of 8 days each for the Chebyshev expansion. For this limited accuracy of few 10^{-4}, the number of adjusted coefficients is similar in both methods (55 vs. 60), but the spline is continuous (and even smooth) everywhere, while the Chebyshev is discontinuous at each subinterval boundaries (here $t = -12, -4, 4, 12$).

The first example in Fig. 7.2 shows a standard comparison when only a relatively low accuracy, slightly better than 10^{-3}, is sought. The Chebyshev representation of degree $d = 10$ over the 5 sub-intervals requires the determination and use of 55 coefficients to evaluate the function at a particular time. I found that one needs about 60 nodes in the spline to achieve similar perfor-

mance over the 40 days, giving a volume of data nearly identical with both methods. However it is conspicuous that the error distributions are quite different. The trademark of the Chebyshev approximation is precisely the evenness in the distribution of error over each sub-interval, even though the typical magnitude could be different from one to the next. Here the oscillations are regular between $\pm 3 \times 10^{-3}$, more or less proportional to $T_{11}(x)$ and crossing the $t-$axis 11 times per sub-interval. The oscillatory feature is more pronounced with spline,showing a strong sub-structure at smaller scale. However the must salient feature is the large increase of the error near the external boundaries, due to the fact the standard spine I use here is not much constrained at the limits. Clamped spline using the values of the derivatives of $f(x)$ at both ends would perform better. In practice this is not a major inconvenient if we have a single spline over a long stretch of time, so that one can restrict its use to a certain fraction, (~ 90 %) inside the full interval.

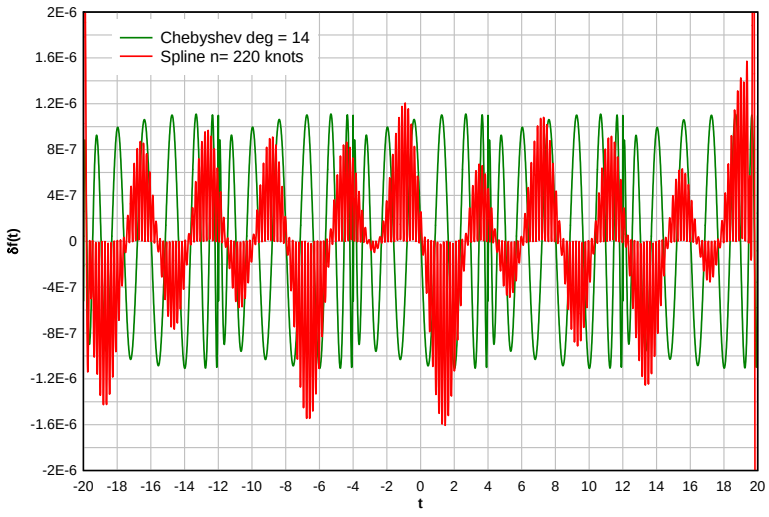


FIGURE 7.3: Comparison of the error in Chebyshev approximation (degree $d = 14$, green curve) to a spline interpolation (220 knots, red curve) for the function defined by (6.1) over a period of 40 days and using 5 subintervals of 8 days each for the Chebyshev expansion. To achieve an accuracy of few 10^{-7}, the number of nodes grows very large and one has four times as many coefficients with the spline than with the Chebyshev expansion. The discontinuity is illustrated in Fig. 7.5

To get a better accuracy with the spline one needs to increase the number of nodes, or equivalently make the interval between two successive nodes smaller. For the Chebyshev, as seen in Sect. 3, one can either use smaller sub-interval or increase the degree (or do both). I play just on the degree to make the comparisons simpler. The accuracy increases very significantly with the degree and using now $d = 14$ leads to an error better than 10^{-6}, just by increasing the comput-

ing cost by 40%. As shown in Fig. 7.3, to achieve similar gain with splines, one must increase considerably the number of nodes, from 60 to 220 to match the Chebyshev performance. This is a general property and a major difference between the two forms of approximation: the error in the Chebyshev approximation decreases very quickly with larger degrees, while the improvement of the spline interpolation is slow as we take more knots on a given interval. While we have a theoretical measure (Sect. 3) for the Chebyshev approximation on well-behaved function, I know of nothing comparable with the splines, but I have not explored the issue as seriously as for the Chebyshev approximation.

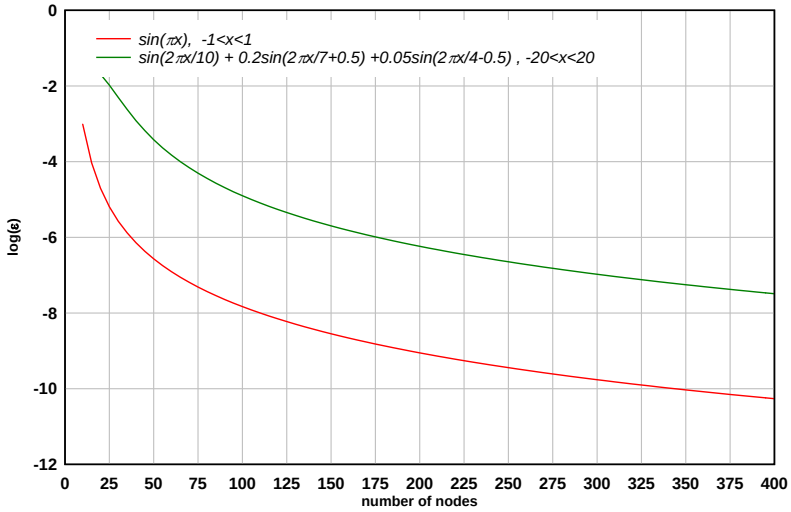


FIGURE 7.4: Accuracy of a spline interpolation for two simple functions with the number of nodes.

A simple numerical illustration is shown in Fig. 7.4 where the accuracy, measured by the average magnitude of the error, is plotted as a function of the number of nodes used on the interpolation interval. The two functions are the simple sine function used in the first section of this note, and the quasi-periodic function over 40 days. To get a good accuracy, as needed for Gaia in the ephemeris representation, the number of nodes becomes very large even for a function without much oscillation over the considered interval, making the recourse to splines not advisable in this case.

Independently of the accuracy a close up view of Fig. 7.5 near the boundary of a sub-interval (here that at $t = 4$) is shown in Fig. 7.5. These boundaries are meaningful only for the piecewise Chebyshev approximation, with a significant discontinuity, while obviously nothing special happens for the spline. The sample points are superposed to a continuous plot to show the jump at $t = 4$ in the Chebyshev approximation. It is of the order of the typical error of the

approximation and in principle should not be a problem in the function evaluation, at least one does not use this approximation to compute first or higher differences for other purposes (numerical differentiation or within a sophisticated integrator). That could cause an instability or in the worst case an overflow during a computation.

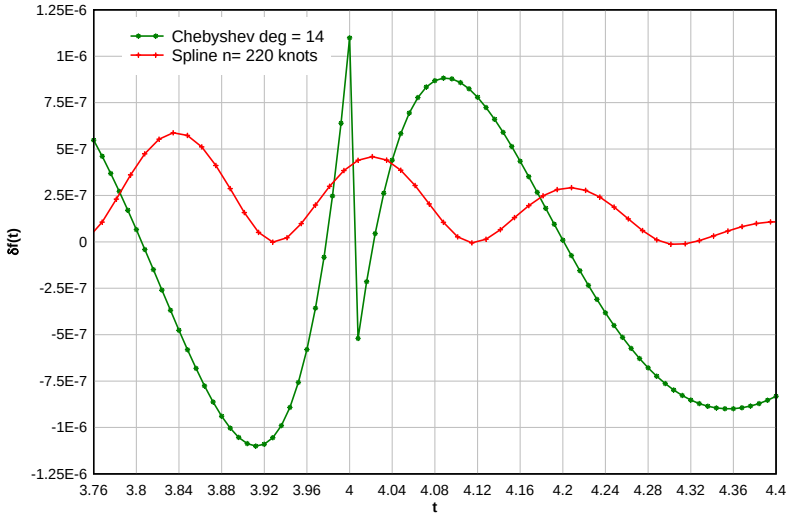


FIGURE 7.5: Close-up of Fig. 7.3 around the discontinuity in the Chebyshev representation at $t = 4$, a boundary between two subintervals. There is no similar boundary for the spline interpolation and no discontinuity.

I have shown in Sect. 4 that one can generate near Chebyshev expansion directly from an interpolation by selecting the interpolating nodes, either at the zeros or the maximums (plus end points) of $T_n(x)$. In the latter case I have also proved that this choice is equivalent to computing the Chebyshev coefficients with a trapezoidal quadrature, avoiding the use of a Lagrange interpolation followed by a conversion of the interpolating polynomials into a Chebyshev form. The result for the same test function is plotted in Fig. 7.6 using $d = 14$ for the degree of the Chebyshev polynomials. One sees by comparison to Fig. 7.3 that the error is typically only larger by a factor 2. This was proved in Sect. 5, Eq. 5.34. This is an acceptable cost knowing that it can be more than absorbed by increasing the degree to 15.

The advantage is seen in Fig. 7.7 with the same close-up as before in the vicinity of a subinterval boundary. The discontinuity in the Chebyshev representation (this not strictly a Chebyshev expansion in the sense of a Fourier expansion in an Hilbert space) has vanished. The discontinuity of the first derivative is hardly visible, but this is just a fortunate circumstance as it can be large.

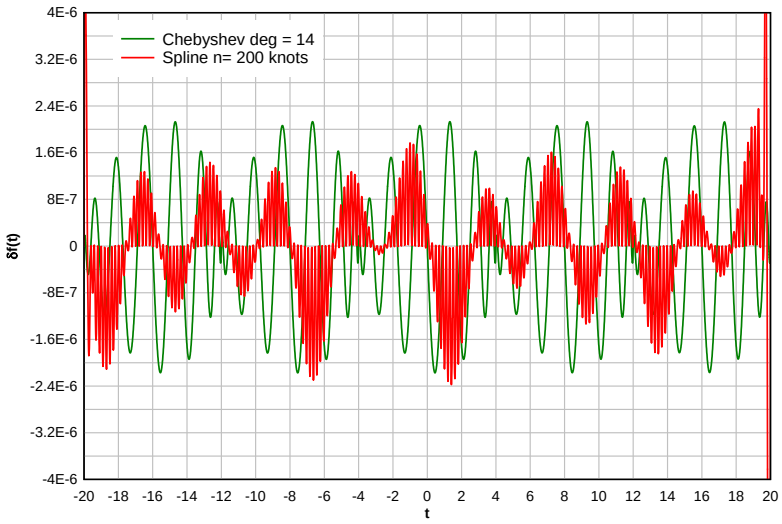


FIGURE 7.6: Comparison of the error in Chebyshev approximation (degree $d = 14$, green curve) computed with the trapezoidal quadrature, to ensure the continuity, to a spline interpolation (200 knots, red curve) for the function defined by (6.1) over a period of 40 days and using 5 subintervals of 8 days each for the Chebyshev expansion. The accuracy is not as good as with the pure Chebyshev expansion of Fig. 7.3, within a factor 2, (it would be 5 times better with $d = 15$) but needs much less coefficients than the spline interpolation with the same accuracy and preserves continuity at the boundary (Fig. 7.7.

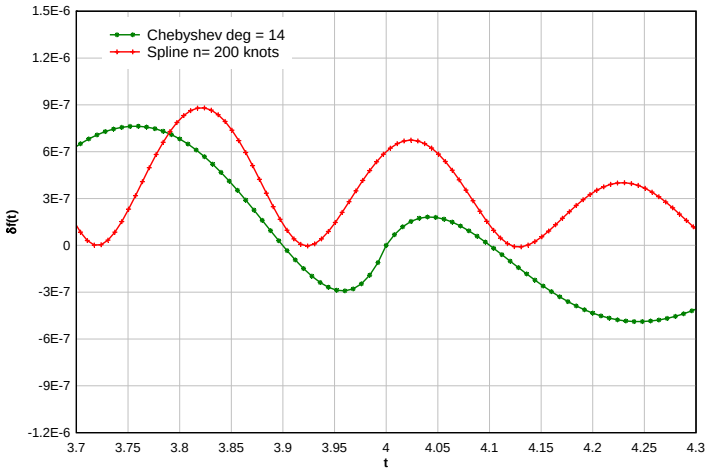


FIGURE 7.7: Close-up of Fig. 7.6 around a boundary of a subinterval of the Chebyshev representation at $t = 4$. Unlike Fig. 7.5 there is no longer discontinuities at the boundary crossings.

8 Conclusions

In this note I have brought into a common framework different ways of approximating a continuous function and expressing the result in the form of a finite expansion on the Chebyshev polynomials.

8.1 Terminology

It would be good to reserve the usage of the term *Chebyshev approximation* to the standard expansion in $\mathcal{L}_2$ associated to the inner product,

$$(f,g) = \int_{-1}^{1} \frac{f(x)g(x)\,dx}{(1-x^2)^{1/2}} \tag{8.1}$$

with the truncated expansion

$$\tilde{f}(x) \approx \phi(x) = {\sum_{j=0}^{k}}' c_j\, T_j(x) \tag{8.2}$$

where the c_j (or their numerical approximation C_j) follow from the least squares minimimisation,

$$\min_{c_0,\cdots,c_k} ||f-\phi||_2 \tag{8.3}$$

The two other approximations based on interpolations at the Chebyshev or Lobatto points, are not stricto sensu true *Chebyshev approximations*, but the difference is very small and perfectly understood. One could call them respectively the *Chebyshev interpolation* and *Lobatto interpolation*, knowing that they are expressed as expansions on the Chebyshev polynomials for convenience.

8.2 Algorithm for the Lobatto interpolation

Since the continuity can be achieved as easily as the pure Chebyshev expansion and without significant loss in accuracy, I would recommend computing all Chebyshev representations of a solar system ephemeris with the Lobatto interpolation using the following algorithm.

Let the function $f(t)$ and the sub-interval of length $2h$ centred at t_0, thus $t_{\text{min}} = t_0 - h$ and $t_{\text{max}} = t_0 + h$. Conversely $t_0 = (t_{\text{min}} + t_{\text{max}})/2$ and $t_0 = (t_{\text{max}} - t_{\text{min}})/2$. The transformations

between the physical variable t and the Chebyshev normalised variable x are given by,

$$x(t) = \frac{t - t_0}{h} \tag{8.4}$$

for the direct transformation and

$$t(x) = t_0 + h\,x \tag{8.5}$$

for the inverse. I take the notation $\tilde{f}(x) = f(t(x))$

Then the coefficients C_j of the expansion to degree k,

$$\tilde{f}(x) \approx \phi(x) = {\sum_{j=0}^{k}}'' C_j\,T_j(x) \tag{8.6}$$

are computed with,

$$C_j = \frac{2}{k}\,{\sum_{i=0}^{k}}''\,\tilde{f}(\zeta_i)\,T_j(\zeta_i),\quad j = 0,\,1,\,\cdots,\,k \tag{8.7}$$

where $T_k(\zeta_i) = \pm 1$, that is to say the ζ_i $(i = 0, \cdots, k)$ are the locations of the extremums of $T_k(x)$ given in (2.10) , including the two end-points.

For the actual publications, one should actually put in the final file the coefficients $\tilde{C}_j$ with

$$\tilde{C}_0 = \frac{C_0}{2} \tag{8.8a}$$

$$\tilde{C}_j = C_j,\ 1 \le j < k \tag{8.8b}$$

$$\tilde{C}_k = \frac{C_k}{2} \tag{8.8c}$$

so that the user will not bother with the halves to evaluate approximations at time t with,

$$f(t) \approx \sum_{i=0}^{k} \tilde{C}_j\,T_j(x(t)) \tag{8.9}$$

8.3 Future work

A comparison between the continuous representation achieved with the interpolation at the Lobatto points and the alternative method proposed by S. Klioner (SK-012) is the obvious next step. In Klioner's proposal, the sampling is regular in t and many points are needed to avoid the equivalent of the Runge effect. By many points, I mean a number several times larger than the degree chosen for the polynomial. This was so designed to use directly the output of a numerical integration naturally with constant step in time. It would be more efficient and would use a much smaller set of data if one could sample the data on the Chebyshev points, with a much regular distribution of the errors over each sub-interval. Normally the use of Lagrange multipliers is not affected by the choice of the sampling and this should work quite as well. The output of a numerical integration can be easily interpolated with central differences between to sample points to any degree of accuracy. Then a comparison with the Lobatto interpolation will be straightforward. This investigation is planned but not yet done and will be reported in a separated note.

9 References

[SK-012], Klioner, S., 2010, *Relativistic time scales and time transformations for Gaia*, GAIA-CA-TN-LO-SK-012, URL http://www.rssd.esa.int/llink/livelink/open/3042112

[GAIA-FM-013], Mignard, F., 2003, *Prototype of the ephemeris of the Earth for GAIA*, GAIA-FM-013, URL http://www.rssd.esa.int/llink/livelink/open/357550

Atkinson, K. E., 1989, An Introduction to Numerical Analysis, 2nd edition, John Wiley & Sons, Inc.

Davis, P. J., 1974, Interpolation and Approximation, Dover Publication.

Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, 1992, Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge, England: Cambridge University Press, pp. 172-178.

Ralston, A. & Rabibowitz, P., 1978, A first course in numerical analysis, Mc Graw Hill.

Rivlin, T. J.,1981, An Introduction to the approximation of functions, Dover Publication.