A COLLOCATION METHOD FOR THE SEQUENTIAL TESTING OF A GAMMA PROCESS

Bruno Buonaguidi and Pietro Muliere

L. Bocconi University, Milan. Italy

Supplementary Material

A Technical proofs of the results in Section 2

A.1 Proof of Proposition 2.1

The expressions (2.13)-(2.15) can be obtained by applying the results of Buonaguidi and Muliere (2013, Sec. 5.2) or can be derived by using Ito's formula and (2.12):

$$
\begin{aligned}
f(\pi_t) =& f(\pi) + \int_0^t f'(\pi_{s^-})\, d\pi_s + \sum_{0\le s\le t} \left(\Delta f(\pi_s) - f'(\pi_{s^-})\Delta\pi_s\right) \\
=& f(\pi) - \int_0^t \log\left(\frac{\alpha_0}{\alpha_1}\right) f'(\pi_{s^-})\pi_{s-}(1-\pi_{s^-})\, ds + \int_0^t \int_0^1 \left[f(\pi_{s^-}+z) - f(\pi_{s^-})\right] \mu^\pi(dz, ds) \\
=& f(\pi) - \int_0^t \log\left(\frac{\alpha_0}{\alpha_1}\right) f'(\pi_{s^-})\pi_{s-}(1-\pi_{s^-})\, ds + \int_0^t \int_0^1 \left[f(\pi_{s^-}+z) - f(\pi_{s^-})\right] v^\pi(dz)\, ds \\
&+ \int_0^t \int_0^1 \left[f(\pi_{s^-}+z) - f(\pi_{s^-})\right] \left(\mu^\pi(dz, ds) - v^\pi(dz)\, ds\right),
\end{aligned}
\tag{A.1}
$$

where μ^π and v^π are the jumping measure and the associated compensator of $(\pi_t)_{t\ge 0}$. From (2.12) one may notice that the magnitude of its jumps is

$$
\Delta\pi_t = \frac{\pi_{t^-}(1-\pi_{t^-})\left(e^{(\alpha_0-\alpha_1)x} - 1\right)}{1+\pi_{t^-}\left(e^{(\alpha_0-\alpha_1)x} - 1\right)},
\tag{A.2}
$$

so that

$$
\pi_{t^-} + \Delta\pi_t = \frac{\pi_{t^-} e^{-\alpha_1 x}}{(1-\pi_{t^-})e^{-\alpha_0 x} + \pi_{t^-} e^{-\alpha_1 x}}.
\tag{A.3}
$$

Hence, the replacement in (A.1) of $(\pi_{s^-} + z)$ with (A.3) and the integration over $(0,\infty)$ with respect to μ^X and its compensator $(1-\pi)v_0 + \pi v_1$, being $v_i(dx) = x^{-1}e^{-\alpha_i x}\mathbf{1}_{(0,\infty)}(dx)$, $i = 0, 1$, complete the proof.

A.2 Proof of Proposition 2.2

Let π_1, $\pi_2 \in [0,1]$ and $\lambda \in [0,1]$. From (2.3), it is immediate to notice that $P_{\lambda\pi_1+(1-\lambda)\pi_2} = \lambda P_{\pi_1} + (1-\lambda) P_{\pi_2}$. Hence,

$$\begin{aligned} V\big(\lambda\pi_1 + (1-\lambda)\pi_2\big) &= \inf_\tau E_{\lambda\pi_1+(1-\lambda)\pi_2}\left[\tau + g_{a,b}(\pi_\tau)\right] \\ &= \inf_\tau \Big\{ \lambda E_{\pi_1}\left[\tau + g_{a,b}(\pi_\tau)\right] + (1-\lambda) E_{\pi_2}\left[\tau + g_{a,b}(\pi_\tau)\right] \Big\} \\ &\geq \lambda \inf_\tau E_{\pi_1}\left[\tau + g_{a,b}(\pi_\tau)\right] + (1-\lambda) \inf_\tau E_{\pi_2}\left[\tau + g_{a,b}(\pi_\tau)\right] \qquad (A.4) \\ &= \lambda V(\pi_1) + (1-\lambda) V(\pi_2). \end{aligned}$$

A.3 Proof of Proposition 2.3

Since on (A, B) we have $V(\pi) < g_{a,b}(\pi)$, for any $\epsilon > 0$ such that $A + \epsilon < c$, it results

$$\frac{V(A+\epsilon) - V(A)}{\epsilon} \leq \frac{a(A+\epsilon) - aA}{\epsilon} = a, \qquad (A.5)$$

so that $V'(A+) \leq a$, where the right-hand derivative exists because of the concavity of $\pi \mapsto V(\pi)$.

In order to show that the reverse inequality holds, fix $\epsilon > 0$ so that $A + \epsilon < c$ and consider the stopping time $\tau^\star_{A+\epsilon}$, that, according to the arguments of Subsection 2.1, is optimal for $V(A+\epsilon)$. We recall that $\tau^\star_{\pi+\epsilon}$ is the first exit time from (A, B) of the process $(\pi_t)_{t\geq 0}$, starting at $\pi_0 = \pi + \epsilon$. Then, from (2.3) and similarly to Gapeev and Peskir (2004), we have

$$\begin{aligned} &V(A+\epsilon) - V(A) \\ &\quad \geq E_{A+\epsilon}\left[\tau^\star_{A+\epsilon} + g_{a,b}(\pi_{\tau^\star_{A+\epsilon}})\right] - E_A\left[\tau^\star_{A+\epsilon} + g_{a,b}(\pi_{\tau^\star_{A+\epsilon}})\right] = \sum_{i=0}^{1} E_i\left[S_i(A+\epsilon) - S_i(A)\right], \qquad (A.6) \end{aligned}$$

where

$$S_i(\pi) = \frac{1 + (-1)^i(1-2\pi)}{2}\left(\tau^\star_{A+\epsilon} + a\frac{\pi e^{Y_{\tau^\star_{A+\epsilon}}}}{1 + \pi(e^{Y_{\tau^\star_{A+\epsilon}}} - 1)} \wedge b\frac{1-\pi}{1 + \pi(e^{Y_{\tau^\star_{A+\epsilon}}} - 1)}\right). \qquad (A.7)$$

Then, according to the mean value theorem, there exist $\xi_i \in (A, A+\epsilon)$, $i = 0, 1$, such that

$$\sum_{i=0}^{1} E_i\left[S_i(A+\epsilon) - S_i(A)\right] = \epsilon \sum_{i=0}^{1} E_i\left[S_i'(\xi_i)\right], \qquad (A.8)$$

being

$$\begin{aligned} S_i'(\pi) = (-1)^{i-1}&\left(\tau^\star_{A+\epsilon} + a\frac{\pi e^{Y_{\tau^\star_{A+\epsilon}}}}{1 + \pi(e^{Y_{\tau^\star_{A+\epsilon}}} - 1)} \wedge b\frac{1-\pi}{1 + \pi(e^{Y_{\tau^\star_{A+\epsilon}}} - 1)}\right) \\ &+ \frac{1 + (-1)^i(1-2\pi)}{2}\left(a\mathbf{1}_{\left\{\pi_{\tau^\star_{A+\epsilon}} < c\right\}} - b\mathbf{1}_{\left\{\pi_{\tau^\star_{A+\epsilon}} > c\right\}}\right) \frac{e^{Y_{\tau^\star_{A+\epsilon}}}}{\left[1 + \pi(e^{Y_{\tau^\star_{A+\epsilon}}} - 1)\right]^2}. \qquad (A.9) \end{aligned}$$

From the definition of $\tau^\star_{\pi+\epsilon}$ and simple calculations, one has

$$\begin{aligned}\tau^\star_{A+\epsilon} &= \inf\{t \geq 0 : \pi_t \notin (A, B),\ \pi_0 = A + \epsilon\} \\ &\leq \inf\left\{t \geq 0 : Y_t \leq \log\left(\frac{A}{1-A}\frac{1-(A+\epsilon)}{A+\epsilon}\right)\right\} =: \gamma_\epsilon. \end{aligned} \tag{A.10}$$

According to Sato (1999, Th. 43.20, p. 323),

$$P_i\left[\lim_{t\downarrow 0} t^{-1} Y_t = -\log\left(\frac{\alpha_0}{\alpha_1}\right)\right] = 1, \quad i = 0, 1, \tag{A.11}$$

meaning that the starting point 0 of $Y = (Y_t)_{t\geq 0}$ is regular for $(-\infty, 0)$ (that is, with probability 1, Y, starting at 0, enters $(-\infty, 0)$ immediately). From (A.10) and (A.11), it results $\gamma_\epsilon \downarrow 0$ P_i-a.s. as $\epsilon \downarrow 0$, $i = 0, 1$. Therefore, $\tau^\star_{A+\epsilon} \downarrow 0$ and $Y_{\tau^\star_{A+\epsilon}} \to 0$ as $\epsilon \downarrow 0$ P_i-a.s., $i = 0, 1$. Hence, from (A.9)

$$S'_i(\xi_i) \to (-1)^{i-1} aA + \frac{1 + (-1)^i(1-2A)}{2} a, \quad P_i\text{-a.s.},\ i = 0, 1,\ \text{as } \epsilon \downarrow 0. \tag{A.12}$$

Since $S'_i(\xi_i) + (-1)^i \tau^\star_{A+\epsilon}$ is obviously bounded, for $i = 0, 1$, from (A.6), (A.8), (A.12), $E_i[\tau^\star_{A+\epsilon}] \to 0$ as $\epsilon \downarrow 0$, $i = 0, 1$, and the bounded convergence theorem we have

$$V'(A+) = \lim_{\epsilon\downarrow 0} \frac{V(A+\epsilon) - V(A)}{\epsilon} \geq \lim_{\epsilon\downarrow 0} \sum_{i=0}^{1} E_i\left[S'_i(\xi_i)\right] = a, \tag{A.13}$$

which, combined with (A.5), completes the proof.

A.4 Proof of Proposition 2.4

Define $f(y) = V(\pi; B)$, with $\pi = e^y/(1+e^y)$; it is not difficult to show that f solves

$$\begin{aligned} f'(y) = &-\frac{1}{\lambda} - \frac{b}{\lambda}\frac{e^{\gamma y}}{1+e^y}\int_{B^o}^{\infty} \frac{e^{-\gamma z}}{z-y}\,dz + f(y)\frac{e^{\gamma y}}{(1+e^y)\lambda}\int_{B^o}^{\infty}\frac{(1+e^z)e^{-\gamma z}}{z-y}\,dz \\ &- \frac{e^{\gamma y}}{(1+e^y)\lambda}\int_y^{B^o} [f(z) - f(y)]\frac{(1+e^z)e^{-\gamma z}}{z-y}\,dz, \quad y^\star \leq y < B^o, \end{aligned} \tag{A.14}$$

$$f(B^o) = \frac{b}{1+e^{B^o}}, \tag{A.15}$$

where $y^\star$ is any arbitrary finite number smaller than B^o, $B^o = \log(B/(1-B))$, $\gamma = \alpha_0/(\alpha_0 - \alpha_1)$ and $\lambda = \log(\alpha_1/\alpha_0)$. The representation (A.14)-(A.15) is equivalent to (2.24)-(2.25), but has the advantage of directly appearing as a linear Volterra integro-differential equation of the second kind (meaning that one limit of integration is variable and the unknown function f also occurs outside the integral). We observe that (A.14) seems to be outside the scope of any existing theory on integro-differential equations, because one has to consider the difference $f(z) - f(y)$ in the last integral (and not just $f(z)$ like in the canonical representation (B.1)), in order to make it finite. This is caused by the lack of integrability of the map $z \mapsto (1+e^z)e^{-\gamma z}/(z-y)$ on (y, B^o), which, in turn, is a consequence of the Lévy measure of a gamma process. Then, we proceed as follows: first we analyze "regular versions" of (A.14)-(A.15), for which the existence and uniqueness of solutions can be proved by resorting to standard theory; then, we verify that the limit of these solutions is indeed a solution of (A.14)-(A.15).

Let $0 < \epsilon \leq 1$ and denote by $f_\epsilon(y)$ the function solving the following "regular" problem:

$$f'_\epsilon(y) = g(y) + h_\epsilon(y) f_\epsilon(y) + \int_y^{B^o} k_\epsilon(y,z) f_\epsilon(z)\, dz, \quad y^\star \leq y < B^o, \tag{A.16}$$

$$f_\epsilon(B^o) = \frac{b}{1+e^{B^o}}, \tag{A.17}$$

where

$$g(y) = -\frac{1}{\lambda} - \frac{b}{\lambda}\frac{e^{\gamma y}}{1+e^y}\int_{B^o}^\infty \frac{e^{-\gamma z}}{z-y}\, dz, \tag{A.18}$$

$$h_\epsilon(y) = \frac{e^{\gamma y}}{(1+e^y)\lambda}\left[\int_y^{B^o} \frac{(1+e^z)e^{-\gamma z}}{(z-y)^{1-\epsilon}}\, dz + \int_{B^o}^\infty \frac{(1+e^z)e^{-\gamma z}}{z-y}\, dz\right], \tag{A.19}$$

$$k_\epsilon(y,z) = -\frac{e^{\gamma y}}{(1+e^y)\lambda}\frac{(1+e^z)e^{-\gamma z}}{(z-y)^{1-\epsilon}}. \tag{A.20}$$

Expressing (A.16)-(A.17) as a system of integral equations

$$w_\epsilon(y) = g(y) + h_\epsilon(y) f_\epsilon(y) + \int_y^{B^o} k_\epsilon(y,z) f_\epsilon(z)\, dz, \tag{A.21}$$

$$f_\epsilon(y) = \frac{b}{1+e^{B^o}} - \int_y^{B^o} w_\epsilon(z)\, dz, \tag{A.22}$$

or, more compactly,

$$F_\epsilon(y) = G_\epsilon(y) + \int_y^{B^o} K_\epsilon(y,z) F_\epsilon(z)\, dz, \tag{A.23}$$

where

$$F_\epsilon(y) = \begin{bmatrix} w_\epsilon(y) \\ f_\epsilon(y) \end{bmatrix}, \quad G_\epsilon(y) = \begin{bmatrix} g(y) + h_\epsilon(y) b/\big(1+e^{B^o}\big) \\ b/\big(1+e^{B^o}\big) \end{bmatrix}, \quad K_\epsilon(y,z) = \begin{bmatrix} -h_\epsilon(y) & k_\epsilon(y,z) \\ -1 & 0 \end{bmatrix}, \tag{A.24}$$

and using the matrix norm $||K_\epsilon(y,z)|| = \max\{h_\epsilon(y) + |k_\epsilon(y,z)|, 1\}$, the following facts are easily verified: i) $G_\epsilon(y)$ is a continuous function of y, in the sense that its components are all continuous; ii) for every continuous vector function s and all $y \leq n_1 \leq n_2 \leq B^o$, $\int_{n_1}^{n_2} K_\epsilon(y,z) s(z)\, dz$ is a continuous function of y; iii) every component of $K_\epsilon(y,z)$ is absolutely integrable with respect to z, for $y^\star \leq y < B^o$; iv) $\exists\ y^\star = y_0 < y_1 < \ldots < y_n = B^o$ such that, for all $i = 0, \ldots, n-1$, $\int_{y_i}^{\min\{y, y_{i+1}\}} ||K_\epsilon(y,z)||\, dz \leq p < 1$, where p is independent of y and i; v) for $y^\star \leq y \leq B^o$, $\lim_{\delta\downarrow 0} \int_{y-\delta}^{y} ||K_\epsilon(y-\delta, z)||\, dz = 0$. Then, according to Linz (1985, Th. 3.2, p. 32), we can conclude that for any $0 < \epsilon \leq 1$, there exists only one continuous solution $F_\epsilon(y)$ to (A.23), that is, the integro-differential equation (A.16)-(A.17) has a unique continuously differentiable solution f_ϵ.

A direct analysis based on the existence and uniqueness of f_ϵ, $0 < \epsilon \leq 1$, shows that $\{f_\epsilon\}$ and $\{f'_\epsilon\}$ are Cauchy sequences and therefore are uniform convergent on $[y^\star, B^o]$. Then

$$f(y) := \lim_{\epsilon\downarrow 0} f_\epsilon(y), \quad f'(y) := \lim_{\epsilon\downarrow 0} f'_\epsilon(y), \quad y^\star \leq y \leq B^o, \tag{A.25}$$

exist and we have that f is continuously differentiable with derivative f'. Further, since

$$\lim_{\epsilon\downarrow 0}\frac{f_\epsilon(z)-f_\epsilon(y)}{(z-y)^{1-\epsilon}}=\frac{f(z)-f(y)}{(z-y)} \quad \text{and} \quad \left|\frac{f_\epsilon(z)-f_\epsilon(y)}{(z-y)^{1-\epsilon}}\right|\le C_y \tag{A.26}$$

for any $z\in[y,B^o]$ and $0<\epsilon\le 1$, where C_y is a constant depending on y, from the bounded convergence theorem we get

$$\begin{aligned} f'(y)=\lim_{\epsilon\downarrow 0}f'_\epsilon(y) &= -\frac{1}{\lambda}-\frac{b}{\lambda}\frac{e^{\gamma y}}{1+e^y}\int_{B^o}^\infty \frac{e^{-\gamma z}}{z-y}\,dz+\lim_{\epsilon\downarrow 0}f_\epsilon(y)\frac{e^{\gamma y}}{(1+e^y)\lambda}\int_{B^o}^\infty\frac{(1+e^z)e^{-\gamma z}}{z-y}\,dz \\ &\quad -\frac{e^{\gamma y}}{(1+e^y)\lambda}\lim_{\epsilon\downarrow 0}\int_y^{B^o}[f_\epsilon(z)-f_\epsilon(y)]\frac{(1+e^z)e^{-\gamma z}}{(z-y)^{1-\epsilon}}\,dz \\ &= -\frac{1}{\lambda}-\frac{b}{\lambda}\frac{e^{\gamma y}}{1+e^y}\int_{B^o}^\infty\frac{e^{-\gamma z}}{z-y}\,dz+f(y)\frac{e^{\gamma y}}{(1+e^y)\lambda}\int_{B^o}^\infty\frac{(1+e^z)e^{-\gamma z}}{z-y}\,dz \\ &\quad -\frac{e^{\gamma y}}{(1+e^y)\lambda}\int_y^{B^o}[f(z)-f(y)]\frac{(1+e^z)e^{-\gamma z}}{z-y}\,dz, \quad y^\star\le y<B^o. \end{aligned} \tag{A.27}$$

Hence, f from (A.25) is a continuously differentiable solution of (A.14)-(A.15), that is, (2.24)-(2.25) admits a continuously differentiable solution $V(\pi;B)$, $\pi\in I_B$. The probabilistic argument provided at the end of the proof of Theorem 2.1 below finally shows that $V(\pi;B)$ is unique.

A.5 Proof of Proposition 2.5

The existence and uniqueness of the map $\pi\mapsto V(\pi;B)$, $\pi\in I_B$, $c<B<1$, has been previously stated. The necessity and sufficiency of (2.27) for having a unique pair $A^\star$ and $B^\star$ solving (2.29), and therefore a unique solution of the free-boundary problem (2.17)-(2.22), arise from the following reasoning.

A direct verification based on the arguments of Section 3 (or the more formal proof given by Peskir and Shiryaev (2000, Remark 2.2, p. 850)) shows that the maps $\pi\mapsto V(\pi;B')$ and $\pi\mapsto V(\pi;B'')$, $B'<B''$, do not intersect on the interval $(0,B']$ (see Figure 3). Condition (2.27) guarantees that for $B>c$, close enough to c, $\pi\mapsto V(\pi;B)$ crosses $\pi\mapsto a\pi$ at some $\pi<c$. Then moving B from c to 1, it is easily seen that there exists a unique pair $A^\star$ and $B^\star$ satisfying (2.29). In other words, there exists a unique pair $A^\star$ and $B^\star$ at which V, provided by (2.28), is consistent with (2.20)-(2.22).

A.6 Proof of Theorem 2.1

The second statement of the theorem is obvious and more arguments can be found in Peskir and Shiryaev (2000, pp. 849-850). According to Buonaguidi and Muliere (2013, Th 5.1, p. 58), for proving the first part of the theorem we only need to check that $(\mathbb{L}V)(\pi)\ge -1$, for $\pi\in[0,1]$, where $\mathbb{L}$ is given in (2.14). By construction, this condition is satisfied on the interval $(A^\star,B^\star)$. For $\pi\in(B^\star,1]$, on which $V(\pi)=b(1-\pi)$, a simple application of the Frullani's integral (2.8) shows that $(\mathbb{L}V)(\pi)=0$. When $\pi=A^\star$, the smooth and continuous fit conditions (2.20) and (2.21) imply $(\mathbb{L}V)(A^\star)=-1$. Finally, one can easily show that $(\mathbb{L}V)(A^\star-)=-1$ that, along with $\partial(\mathbb{L}V)(\pi)/\partial\pi\le 0$ for $\pi\in[0,A^\star)$, completes the proof.

We remark that the following probabilistic argument can be used to prove that for any $B>c$ the map $\pi\mapsto V(\pi;B)$, $\pi\in I_B$, solving (2.24)-(2.25), is unique. Let $g(\pi)=(m\pi+q)\wedge b(1-\pi)$, where $\pi\mapsto m\pi+q$ is the line hitting smoothly $\pi\mapsto V(\pi,B)$ at some $Z<B$. Consider now the optimal stopping problem (2.6)

with $g(\pi)$ in place of $g_{a,b}(\pi)$ and denote by $V(\pi)$ the correspondent value function. Define $V^\star(\pi) = V(\pi; B)$, for $\pi \in (Z, B)$, being $V(\pi; B)$ a solution to (2.24)-(2.25), and $V^\star(\pi) = g(\pi)$, for $\pi \in [0, Z] \cup [B, 1]$. Then, the same arguments of Theorem 2.1 imply that $V(\pi) = V^\star(\pi)$, for $\pi \in [0, 1]$. Since Z is arbitrary, the claim is verified.

B Preliminaries on the collocation method

In Section 3 a numerical scheme, aiming at computing the solution of the free-boundary problem characterizing the sequential testing of a gamma process, is described. Here, we introduce the basic elements on the collocation method and Chebyshev polynomials which our algorithm relies on.

B.1 Collocation method for a linear Volterra integro-differential Equation

Let $\mathbb{T}$ be a linear Volterra integro-differential operator acting on a function f belonging to its domain of definition as

$$(\mathbb{T}f)(x) = f'(x) - g(x) - h(x)f(x) - \int_A^x k(x,z)f(z)\,dz, \quad x \in I = [A, B] \subset \mathbb{R}, \tag{B.1}$$

where $g(x)$, $h(x)$ and $k(x,z)$, $x \in I$ and $A \leq z \leq x$, are known functions. Consider now the functional equation

$$(\mathbb{T}f)(x) = 0, \tag{B.2}$$

along with the boundary condition

$$f(A) = p, \tag{B.3}$$

where p is a fixed number. It is assumed that the boundary value problem (B.2)-(B.3) admits a unique solution f on I that we want to determine. Often this task cannot be analytically accomplished, so that we need numerical techniques allowing us to approximate f as accurately as desired: one of them is the so called collocation method (see, for example, Brunner (2004) or Kress (1998, Sec. 12.4)).

Let us briefly explain its main idea. Let $\Phi = \{\phi_i\}_{i \geq 0}$ be a known basis for f and denote by f_n an approximation of f obtained as linear combination of the first $n+1$ basis functions:

$$f(x) \approx f_n(x) = \sum_{i=0}^{n} w_i \phi_i(x), \quad x \in I, \tag{B.4}$$

so that

$$f'(x) \approx f_n'(x) = \sum_{i=0}^{n} w_i \phi_i'(x), \quad x \in I. \tag{B.5}$$

Choosing n points, known as collocation nodes, $x_i \in I$, $i = 1, .., n$, the problem (B.2)-(B.3) boils down to computing the coefficients w_i by solving the following system of $n+1$ linear equations:

$$(\mathbb{T}f_n)(x_i) = 0, \quad i = 1, .., n, \tag{B.6}$$

$$f_n(A) = p. \tag{B.7}$$

Two problems naturally arise: the choice of an appropriate basis for f and of the truncation limit n.

B.2 The Chebyshev polynomials

In addition to the uniqueness of f for the problem (B.2)-(B.3), assume that f is continuous on I. Then, according to the Waierstrass approximation theorem, f can be uniformly approximated on I by polynomials. One could be tempted to use as Φ the family $\{x^i\}_{i\geq 0}$: its drawback is the lack of the orthogonality property.

Let us recall that a family of functions $\{\psi_i\}_{i\geq 0}$ is said to be orthogonal on I with respect to the weighting function $\eta(x)$ if

$$\int_I \psi_i(x)\psi_j(x)\eta(x)\,dx = \begin{cases} 0, & i \neq j \\ \lambda_j, & i = j \end{cases}. \tag{B.8}$$

The idea is that the information set of an element of a family of orthogonal functions does not overlap with the one expressed by another member of the family. Therefore, if we choose as basis for f a family of orthogonal polynomials, the performances in the numerical approximation of f are improved, due to a better identification of the coefficients w_i in (B.4).

A well known family of orthogonal polynomials is the family of Chebyshev polynomials: their detailed description can be found in Hamming (1986, Sec. 2.28 and 2.29) and Lanczos (1988, Chap. 7); here, we illustrate their main properties, which explain why they represent one of the most important family of polynomials (and, maybe, the most important one) in approximation theory.

The Chebyshev polynomials $\{T_i\}_{i\geq 0}$ are defined by

$$T_n(x) = \cos[n(\arccos(x))], \quad n \geq 0, \quad x \in [-1, 1]. \tag{B.9}$$

The trigonometric identity

$$\cos(n+1)\theta + \cos(n-1)\theta = 2\cos\theta\cos n\theta \tag{B.10}$$

and the substitution $\theta = \arccos(x)$ in (B.10) lead to the recurrence relationship

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x), \quad n \geq 2. \tag{B.11}$$

Since $T_0(x) = 1$ and $T_1(x) = x$, $x \in [-1, 1]$, from (B.11) it is easily seen that $\{T_i\}_{i\geq 0}$ is a family of polynomials. It presents some remarkable features: 1) Chebyshev polynomials are orthogonal on $[-1, 1]$ with respect to the weighting function $\eta(x) = (1 - x^2)^{-1/2}$:

$$\int_{-1}^{1} \frac{T_m(x)T_n(x)}{\sqrt{1-x^2}}\,dx = \int_0^{\pi} \cos m\theta \cos n\theta \, d\theta = \begin{cases} 0, & m \neq n \\ \frac{\pi}{2}, & m = n \neq 0 \\ \pi, & m = n = 0 \end{cases} ; \tag{B.12}$$

2) the zeros of the n-th degree polynomial T_n are given by

$$x_j = \cos\left(\left(j - \frac{1}{2}\right)\frac{\pi}{n}\right), \quad j = 1, .., n; \tag{B.13}$$

3) for $n \geq 0$, derivatives are easy to compute; for instance:

$$T_n'(x) = \frac{n\sin[n\arccos(x)]}{\sin[\arccos(x)]}, \quad T_n''(x) = \frac{nx\sin[n\arccos(x)]}{\left(\sin[\arccos(x)]\right)^3} - \frac{n^2T_n(x)}{\left(\sin[\arccos(x)]\right)^2}; \tag{B.14}$$

4) the shifted Chebyshev polynomials on the interval $I = [A, B]$, $\{T_i^I\}_{i\geq 0}$, along with their first and second derivatives, $\{T_i^{'I}\}_{i\geq 0}$ and $\{T_i^{''I}\}_{i\geq 0}$, are given, for $n \geq 0$ and $x \in I$, by

$$T_n^I(x) = T_n\left(2\frac{x-A}{B-A} - 1\right), \tag{B.15}$$

$$T_n^{'I}(x) = \frac{2}{B-A}T_n'\left(2\frac{x-A}{B-A} - 1\right), \quad T_n^{''I}(x) = \frac{4}{(B-A)^2}T_n''\left(2\frac{x-A}{B-A} - 1\right); \tag{B.16}$$

5) Chebyshev expansions are usually one of the most rapidly convergent expansions for functions (see, e.g., Boyd and Petschek (2014)).

Properties 1-5 appropriately justify the use of Chebyshev polynomials as basis for f; in particular, according to the fifth property, which does not hold only in isolated cases, "low degree" polynomials often lead to satisfactory approximations; in turn, this reflects in a saving of time during numerical computations.

B.3 Accuracy of the Solution

Once a basis for the function f in the problem (B.2)-(B.3) has been chosen, we should determine the length n of the expansion in (B.4).

The truncated series (B.4), whose coefficients are obtained as solution of (B.6)-(B.7), approximately solves (B.2), in the sense that if we replace (B.4) and (B.5) in (B.2), then $(\mathbb{T}f_n)(x) \approx 0$, $x \in I$. This suggests we could increase n until

$$\sup_{x\in I} |(\mathbb{T}f_n)(x)| < \epsilon \tag{B.17}$$

for a fixed $\epsilon > 0$. Of course, since it is not practically possible to evaluate $(\mathbb{T}f_n)(x)$ for any $x \in I$, we can consider a set of equally spaced nodes in I (not the collocation nodes, where $(\mathbb{T}f_n)(x)$ is almost exactly zero) to assess the quality of the computed solution. Alternatively, defining

$$\delta_n = \sup_{x\in I} |f_n(x) - f_{n-1}(x)|, \quad n \geq 1, \tag{B.18}$$

we might increase n until $\delta_n < \delta$, for a specified $\delta > 0$.

We recall that when f is approximated by $f_n = \sum_{i=0}^n w_i T_i^I$, the distance $\sup_{x\in I} |f(x) - f_n(x)|$ is minimized if the collocation nodes are the zeros of T_n^I given by

$$x_j^I = \frac{(B-A)(x_j+1)}{2} + A, \quad j = 1, .., n, \tag{B.19}$$

where x_j is given in (B.13). We observe that the zeros of T_n^I can be used as collocation nodes only if I is known: this does not occur in free-boundary problems, where A and B must be determined. This problem is handled in Section 3.