Wavelets for adaptive solution of boundary value problems *

J. B. Caillau[†] J. Noailles[‡]

Abstract

An adaptive discretization technique for boundary value problems is proposed; the mesh computation essentially involves a wavelet analysis. The process is used to enhance usual methods such as finite differences for linear BVPs or multiple shooting for nonlinear ones. Comparisons with classical non-adaptive Chebyshev spectral or peuso-spectral methods are drawn, in particular in the case of a BVP issuing from nonlinear optimal control.

Key words: boundary value problems, adaptive discretization, wavelet analysis, multiple shooting and finite difference methods, nonlinear optimal control.

AMS subject classifications: 65L10, 65L50, 42C99, 49M05.

1 Introduction

Among the classical methods for the numerical solution of boundary value problems with ordinary differential equations, we can distinguish several parametrization levels of the underlying time discretizations: while spectral or pseudo-spectral techniques [1, 8] are only parametrized by the number of points used for the quadratures, finite difference or more generally multiple shooting approaches depend on up to three parameters: the number of shooting points, the number of steps on each shooting interval, and the number of points used for the integration formula at each step. Modern integrators control the step as well as the quadrature degree, but usually do not provide a mechanism to choose the shooting points themselves. In this regard, we propose an adaptive discretization technique based on a wavelet analysis (see also [9]) that takes advantage of the information currently available on the problem being solved to define dynamically a finite difference or shooting mesh. This approach is used to improve the well-known multiple shooting technique and compares quite favourably with polynomial methods.

The paper is organized as follows. We first define in §2 our wavelet-based technique of mesh computation. Then, we use it in §3 to enhance a finite difference scheme for linear BVPs. Two examples are studied which demonstrate serious advantages not only over direct non-adaptive finite differences but also on a yet robust Chebyshev pseudo-spectral method. Finally, in §4, we deal with the nonlinear BVP coming from an optimal control problem, namely a minimum time orbit transfer in celestial mechanics [11]. Again, coupled with multiple shooting, our procedure proves to be numerically relevant. In particular, it turns to be much more efficient than a modification of the Chebyshev spectral method introduced in [12].

2 Adaptive discretization

Given a general BVP on the fixed interval $[t_0, t_f]$

(1) $$\dot{y} = \xi(t, y)$$

(2) $$b(y(t_0), y(t_f)) = 0$$

*This research was supported in part by the French Ministry for Higher Education and Research (grant #20INP96).

†ENSEEIHT-IRIT, UMR CNRS 5505, 2 rue Camichel, 31071 Toulouse, France, caillau@enseeiht.fr

‡Same address, jnoaille@enseeiht.fr

with ξ and b smooth[1] on appropriate open subsets of $\mathbf{R}\times\mathbf{R}^n$ and $\mathbf{R}^{2n}$ respectively ($b:\mathbf{R}^{2n}\to\mathbf{R}^n$), we suppose the existence of a solution $\bar{y}$ in $C_n^\infty([t_0,t_f])$. Then, subdividing $[t_0,t_f]$ into N subintervals $[t_i,t_{i+1}],\ i=0,\dots,N-1$, one can find an open neighbourhood V_i of each $\bar{y}(t_i)$ in $\mathbf{R}^n$ such that the smooth maximal flow [2] φ_i of $\dot{y}=\xi(t,y)$ with initial condition at t_i is defined on an open subset containing $[t_i,t_{i+1}]\times V_i$. Accordingly, the *multiple shooting* function S associated with (1-2) maps $V_0\times\cdots\times V_{N-1}\times\mathbf{R}^n$ into $\mathbf{R}^{(N+1)n}$ and writes:

$$S(y_0,\dots,y_N)=\begin{bmatrix}\varphi_0(t_1,y_0)-y_1\\ \vdots\\ \varphi_{N-1}(t_N,y_{N-1})-y_N\\ b(y_0,y_N)\end{bmatrix}$$

It is then equivalent to solve (1-2) or to find a zero of S. The numerical evaluation of S is done through some approximation, for instance a *one-step* method (e.g. Runge-Kutta):

(3) $$\varphi_i(t_{i+1},y_i)\simeq y_{i,p},\ i=0,\dots,N-1$$

(4) $$y_{i,0}=y_i,\ y_{i,j+1}=y_{ij}+\Phi(t_{ij},y_{ij},h_{ij}),\ j=0,\dots,p-1$$

(5) $$\Phi(t,y,h)=\sum_{k=1}^{r}\beta_k\xi(t^k,y^k),\ t^k=t+\gamma_k h,\ y^k=y+h\sum_{k=1}^{r}\alpha_{kl}\xi(t^l,y^l),\ k=1,\dots,r$$

We have N shooting points, p (inner) steps and a degree r quadrature formula. Now, these three parameters provide adaptativity to the approximation (3-5); indeed, nowadays integrators automatically adjust the number of steps plus, sometimes, the integration formula degree, so as to control the local error. Furthermore, we propose to compute *well-adapted* shooting points. Before doing this, we just notice that the above-mentionned algorithm can be particularized in several ones: *single shooting*, by taking $N=1$; *finite differences*, by taking $p=1$ (finite differences of degree r, see §3); *spectral* or *pseudo-spectral methods*, by taking N and p equal to 1 (see §3 and 4).

Aiming at computing an adaptive mesh of shooting points, we choose an orthogonal wavelet basis [10] of $\mathrm{L}^2(\mathbf{R})$ $(\psi_{jk})_{jk\in\mathbf{Z}}$, where the mother wavelet ψ is *almost* in the Schwartz class $\mathcal{S}(\mathbf{R})$ of smooth and rapidly decreasing functions (for instance, derivatives of ψ up to a fixed rank decrease faster than any rational function). As a consequence, ψ and its Fourier transform $\widehat{\psi}$ are well-localized (time-frequency localization), and the jk-th wavelet coefficient of a function f in $\mathrm{L}^2(\mathbf{R})$

$$d_{jk}=(f|\psi_{jk})=(\widehat{f}|\widehat{\psi_{jk}})$$

gives us precise information on the behaviour of f around the time point $t=2^{-j}k$ at the frequency 2^j (*local* frequency), since $\psi_{jk}=2^{\frac{j}{2}}\psi(2^jt-k)$ (in turn, $\widehat{\psi_{jk}}$ contains essentially frequencies proportional to 2^j). Now, so as to discretize adaptively f, it is possible to detect its local singularities where a fine mesh is needed: whenever $|d_{jk}|$ is higher than a fixed threshold, discretization points (shooting points, in our case) are added around the abscissa defined by k on the grid of resolution j. This technique is parametrized by:

- *the type of wavelets used*: depending on the family employed (more or less regular, periodic, on the interval...), the discretization changes;
- *the minimum and maximum resolutions*: they depend on the size of details one wants to focus on;
- *the compression ratio*: only a fixed proportion of the entire maximum resolution uniform grid points are kept;
- *the kind of computation*: one can either make a *global computation* of the grid (at all resolutions), or *refine* it by increasing step by step the resolution; moreover, in the case of vector valued functions, one can either *superimpose* the results from each dimension, or *select* the points associated with the biggest coefficients among all dimensions.

The inherent paradox of this method is that, so as to compute a grid which would be perfectly adapted to the problem, one has to know the solution itself. Now, though in practice this is certainly not the case, the iterative process used to solve the multiple shooting equation brings us partial but more and more precise information on the solution: by means of suitable coupling between the iterative procedure and the discretization computation, it is possible to take advantage of this knowledge to construct progressively an adapted mesh. Several levels of connection can be considered:

[1] By *smooth* we mean indefinitely differentiable.

- *with nonlinear iteration*: the Newton-like current iterate is analysed and thus provides a discretization for the next step (see §4);
- *with solving*: the problem is solved repeatedly, first on a coarse grid, and then on finer and finer ones (see §3);
- *with continuation*: as shall be done in §4, it is often desirable to solve a sequence of *mild* problems rather than a single intricate one; then again, the information provided by this continuation procedure can be employed to improve the adaptativity of the discretization.

The approach is first illustrated in the linear case.

3 Linear boundary value problems

We consider a linear BVP

$$\dot{y} = A(t)y + b(t) \tag{6}$$

$$C_0 y(t_0) + C_f y(t_f) = d \tag{7}$$

with linear boundary conditions. We will study two examples, both being constructed in the same way: $A(t)$ is the matrix

$$A(t) = \begin{bmatrix} 1 & -t \\ t & 1 \end{bmatrix}$$

and $b(t)$ is chosen in order that $y = (f, \dot{f})$ be solution of (6-7) (the boundary conditions are $y_1(t_0) = f(t_0),\ y_2(t_f) = \dot{f}(t_f)$). In the first case, f is equal to a pulse function θ_ε, $\varepsilon > 0$ small, where $\theta_\varepsilon = 1/\varepsilon\ \theta(t/\varepsilon)$ with θ the smooth compactly supported function equal to $\exp(1/(|t|^2 - 1))$ on $]-1, 1[$ and 0 anywhere else (clearly, θ_ε tends to the Dirac measure in the *distributions sense* when $\varepsilon \to 0$). For our second example, we take $f = \omega_\varepsilon$, ε strictly positive and small, $\omega_\varepsilon = \sin(1/(t + \varepsilon))$ (function with more than 150 oscillations for $\varepsilon = 0.002$).

Because of the linearity, the flow is affine,

$$\varphi_i(t, y) = Y_i(t)y + \int_{t_i}^{t} Y_i^{-1}(s)b(s)ds$$

with Y_i the fundamental solution of $\dot{y} = A(t)y,\ Y(t_i) = I$, and the shooting equation boils down to a *sparse* linear system (two block diagonals plus the boundary conditions), solved by sparse LU factorization. The approximation is done by finite differences of Gauss type and order $r = 2$ or 3 with *parameter condensation* [1]. The discretization computation is coupled with solving, that is we solve the problem repeatedly on grids that are refined between two iterations: points are added at resolution $j + 1$ around the points related to the highest wavelet coefficients of the current approximation of the solution on the coarser grid of resolution j, with j between 5 and 16. The grid is the *superimposition* of the grids computed on each dimension, and the compression ratio ranges from 2% to 9%. The compactly supported or symmetric wavelets are taken from [3]. The approach is compared with an obvious modification [4, 5] for linear BVPs of the robust pseudo-spectral Chebyshev method of [7] (which uses the non-adaptive grid of Chebyshev polynomials roots).

The results for the first problem (pulse solution) are summarized in table 1 (and figure 1): for $\varepsilon = 0.1$, our method is as precise as the pseudo-spectral one, but more than 300 times faster. Moreover, the intermediate results generated on the adapted sparse non-uniform grids are also as accurate as the ones computed on the full corresponding uniform meshes, and the method is 4 times faster than the computation by finite differences on the only full grid at the finest resolution. For $\varepsilon = 0.01$, the pseudo-spectral method is useless, whereas the computation on the successive adapted meshes is more 13 times faster than the computation only on the finest uniform one. The picture is quite the same on our second example (see also table 1 and figure 1). The computation on the full uniform grid is not provided, since it would require a $N = 2^{16} = 65536$ point mesh. Next section is devoted to the nonlinear case.

4 Nonlinear boundary value problems

The nonlinear BVP under consideration arises when applying the maximum principle to the following optimal control problem: we want to transfer in *minimum time* a satellite from a low orbit to a high geostationnary one [11] (this

Table 1: Numerical results for the two examples (PS=Pseudo-Spectral (Chebyshev), UFD=Uniform Finite Differences, AFD=Adaptive Finite Differences). The Δ_{y_i}'s are absolute errors for the first example, relative errors for the second.

	Example 1, $\varepsilon = 0.1$			Example 1, $\varepsilon = 0.01$			Example 2, $\varepsilon = 0.002$	
	PS	UFD	AFD	PS	UFD	AFD	PS	AFD
N	2000	1024	183	2000	8192	435	2000	7311
T_{exec} (s)	3880	40	10	3720	619	45	3955	509
$\Delta_{y_1(t_0)}$	0.00041	0.00044	0.00042	1710	0.076	0.076	>100%	3%
$\Delta_{y_2(t_f)}$	0.0016	0.0017	0.0017	6930	0.31	0.31	>100%	3%

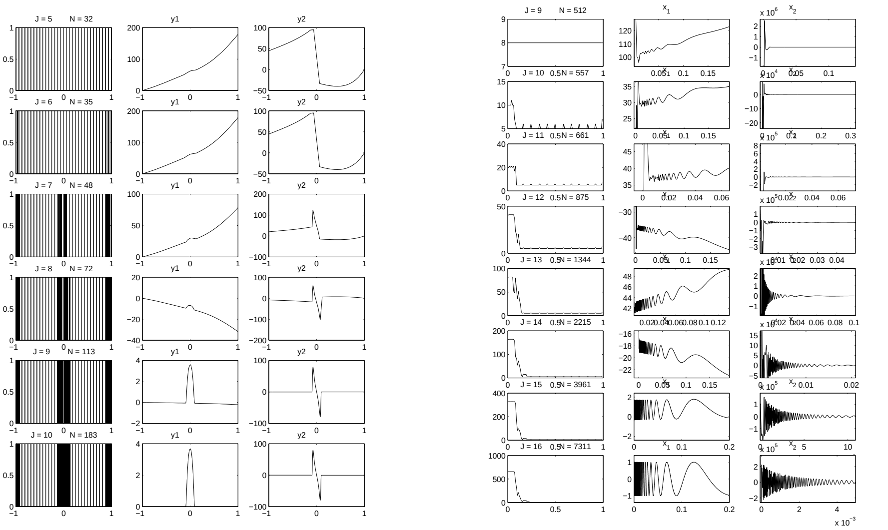

Figure 1: Adaptive finite differences for the first example ($\varepsilon = 0.1$), left, and for the second ($\varepsilon = 0.002$), right (detail). In the latter case, the final sparse grid has 7311 points (against $2^{16} = 65536$ for the uniform full one), and is not represented; rather, the repartitions of the meshes points are given. In both cases, the grid points are concentrated in neighbourhoods of the singularities.

defines the end-point conditions). The dynamics is affine in the control, $\dot{x} = f_0(x) + u_1 f_1(x) + u_2 f_2(x)$, with vector fields defined on a suitable submanifold of $\mathbf{R}^4$ by

$$f_0 = \sqrt{\mu_0/P} \begin{bmatrix} 0 \\ 0 \\ 0 \\ W^2/P \end{bmatrix}, \ f_1 = \sqrt{P/\mu_0} \begin{bmatrix} 0 \\ \sin L \\ -\cos L \\ 0 \end{bmatrix}, \ f_2 = \sqrt{P/\mu_0} \begin{bmatrix} 2P/W \\ \cos L + (e_x + \cos L)/W \\ \sin L + (e_y + \sin L)/W \\ 0 \end{bmatrix}$$

(the parameters $(P, e_x, e_y, L) = x$ describe the osculating ellipse). There is also a constraint on the control, $|u| \leq F_{max}/M$ (maximum thrust allowed for a mass M satellite). Because of the special structure of the Lie algebra generated by f_0, f_1 and f_2, the system turns to be controllable, from where the existence of an optimal control follows. Besides, on the basis of the geometric study of [6], we will assume that the control is smoothly defined everywhere by $u = -F_{max}/M \ (H_1, H_2)/|(H_1, H_2)|$ where $H_i = (p|f_i(x))$, p adjoint state (*no commutation* assumption). Together with transversality conditions, the last relation defines a nonlinear BVP in $y = (x, p)$ in the form (1-2) (the unknown transfer time being treated as an additional state variable, $\dot{t}_f = 0$).

The problem is solved by multiple shooting, with an adaptive computation of shooting points using the wavelets on the interval of [3] and a resolution j ranging from 4 to 13. The computation of the $N = 10$ shooting points is *global* (the whole grid is recomputed at each step) and uses a *selection* (not a superimposition) procedure among all dimensions. The first coupling considered is with the nonlinear Newton iteration: the first iterations are done with single shooting until a prescribed tolerance of the nonlinear solver is reached, then the approximation of the solution is used to compute an adapted mesh (see figure 2). This approach, however, proves to be not very efficient since it becomes compulsary to do almost all the iterations with single shooting in order to obtain convergence. For this reason, we appeal to a second kind of coupling, namely coupling with continuation. Indeed, our problem is parametrized by the maximum thrust, and the results for the current value F^c_{max} can be used to initialize the solving for a lower thrust F^+_{max}. Here, we also employ them to determine adaptively the shooting points. The results are compared with those given by the Chebyshev spectral method for optimal control problems of [12], enhanced with domain decomposition.

Whereas it takes several hours to get a result with the spectral method (e.g. for 9 Newtons, see figure 3), the computation with adaptive multiple shooting does not require more than one minute for the same thrusts (see table 2). Moreover, the adaptive choice of the shooting meshes improves the convergence of the algorithm since direct multiple shooting with $N = 10$ shooting points uniformly distributed diverges below $F_{max} = 20$ Newtons. An example of discretization is given in figure 3.

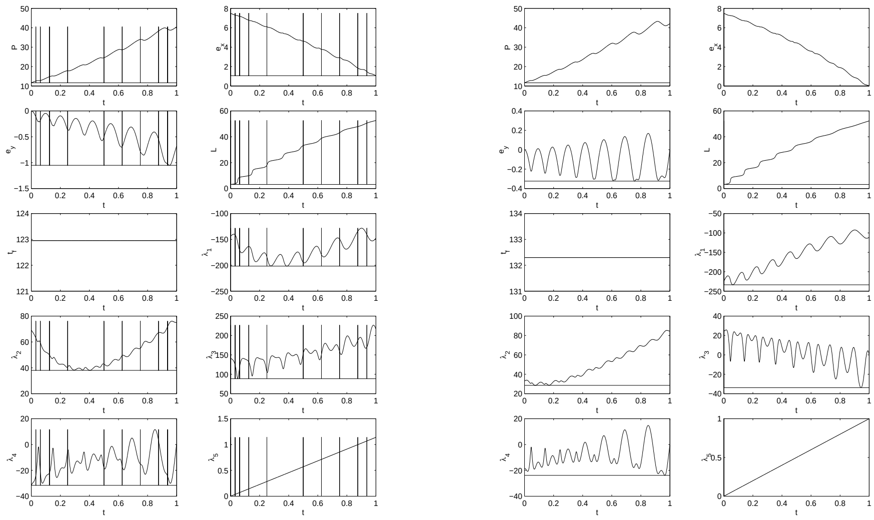

Figure 2: Coupling with nonlinear iteration, $F_{max} = 7$ Newtons. Left, the approximation of the solution provided by the first single shooting iterations is adaptively discretized thanks to the wavelet analysis. The result obtained by finishing the computation with multiple shooting on this grid is on the right.

Table 2: Minimum transfer times and execution times obtained with adaptive multiple shooting. The choice of thrusts for the continuation is heuristic.

F_{max} (Newtons)	$\bar{t}_f$ (Hours)	T_{exec} (Seconds)
30	32.56	3.37
24	35.93	3.66
17	51.51	6.17
12	73.27	19.03
9	100.8	19.78

5 Conclusion

We have illustrated on several examples of BVPs, linear and nonlinear, how it is possible to take advantadge of the intrinsic adaptativity of multiple shooting-like methods thanks to a suitable wavelet analysis. In particular, the results obtained on the sparse non-uniform generated grids are as precise as those obtained on the full uniform ones. Over and above, the amount of computation is drastically decreased, and the convergence is even improved.

It is our belief that this kind of technique, which allows various coupling levels with the associated solving procedure, can be successfully reused to enhance numerous existing non-adaptive algorithms. The developed software is available on request to the authors.

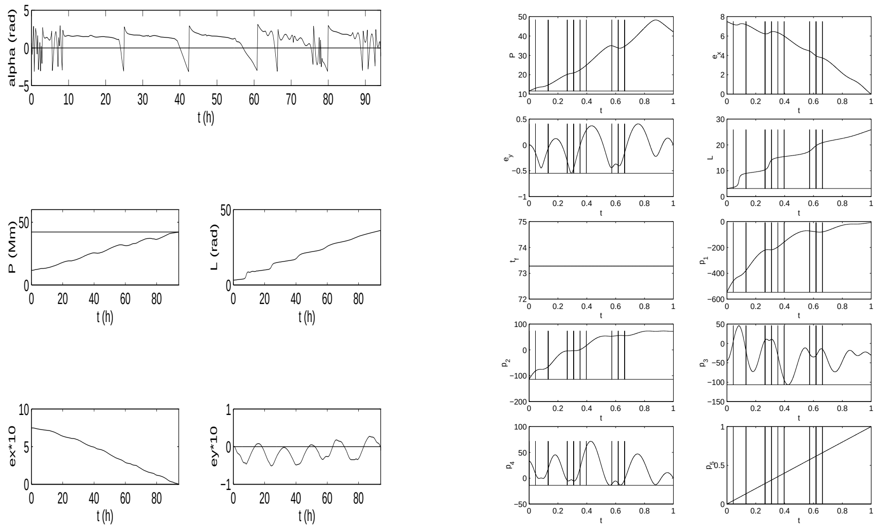

Figure 3: On the left, the solution (control and state) for $F_{max} = 9$ Newtons given by the Chebyshev spectral method (with 12 subdomains and degree 8 polynomials on each for the state and the control [4]); the control sought is in fact the angle α of the command (which is always of maximum modulus by virtue of Pontryagin maximum principle). On the right, the solution (state and ajoint state, with t_f as a new state variable) for $F_{max} = 12$ Newtons. The grid, computed on the result for 17 Newtons, matches fast variations of the state at the perigee ($L \equiv 0(2\pi)$).

References

[1] U. M. Ascher, R. M. M. Mattheij, and R. D. Russel. *Numerical solution of boundary value problems for differential equations.* Prentice Hall, 1988.

[2] M. Berger and B. Gostiaux. *Géométrie différentielle.* Armand-Colin, Paris, 1972.

[3] J. Buckheit, S. Chen, D. Donoho, I. Johnstone, and J. Scargle. WaveLab reference manual. Technical report, Stanford University, 1995.

[4] J. B. Caillau, J. Gergaud, and J. Noailles. Trajectoires optimales à poussée continue (in french). Technical Report R & T A3006, CNES (French Space Agency) / ENSEEIHT-IRIT, July 1998.

[5] J. B. Caillau and J. Noailles. Résolution adaptative de problèmes aux deux bouts linéaires (in french). Séminaire CESAME, Université Catholique de Louvain-la-Neuve, Belgium, June 1998.

[6] J. B. Caillau and J. Noailles. Time optimal orbit transfer. Nonlinear Sciences on the Border of Milleniums, conference dedicated to the 275th anniversary of Russian Academy of Sciences, NS'1999, Saint-Petersburg University, Russia, June 1999.

[7] T. M. El-Gindy, H. M. El-Hawary, M. S. Salim, and M. El-Kady. A Chebyshev approximation for solving optimal control problems. *Computers Math. Applic.*, 29(6):35–45, June 1994.

[8] D. Funaro. *Polynomial approximation of differential equations.* Springer-Verlag, New-York, 1992.

[9] L. Jameson. On the wavelet based differentiation matrix. ICASE report 93-95, NASA Langley Research Center, 1996.

[10] Y. Meyer. *Ondelettes et opérateurs*, volume 1. Hermann, Paris, 1990.

[11] J. Noailles and T. C. Le. Contrôle en temps minimal et transfert orbital à faible poussée (in french). *Équations aux dérivées partielles et applications,* articles in honour of J. L. Lions for his 70th birthday, pages 705–724, Gauthier-Villars, 1998.

[12] J. Vlassenbroeck and R. Van Dooren. A Chebyshev technique for solving nonlinear optimal control problems. *IEEE transactions on automatic control*, 33(4):333–340, April 1988.