Electronic Transactions on Numerical Analysis.
Volume 44, pp. 281–288, 2015.

ISSN 1068–9613.

ETNA
Kent State University
http://etna.math.kent.edu

FAST ALGORITHMS FOR SPECTRAL DIFFERENTIATION MATRICES*

JARED L. AURENTZ†

Abstract. Recently Olver and Townsend presented a fast spectral method that relies on bases of ultraspherical polynomials to give differentiation matrices that are almost banded. The almost-banded structure allowed them to develop efficient algorithms for solving certain discretized systems in linear time. We show that one can also design fast algorithms for standard spectral methods because the underlying matrices, though dense, have the same rank structure as those of Olver and Townsend.

Key words. spectral methods, rank-structured matrices

AMS subject classifications. 65N35, 33C45, 65F05

1. Introduction. In [13] Olver and Townsend presented a fast spectral method for solving linear differential equations using bases of ultraspherical polynomials, which has subsequently been exploited in Chebfun [6] and ApproxFun [12]. The "fast" in these methods appears to come from the fact that the discretized systems involve *almost-banded* matrices, making it possible to solve the systems in linear time via structured QR factorizations.[1,2] More standard bases, such as Chebyshev bases, give dense matrices that appear to require more expensive linear algebra. Here we show that, although standard spectral matrices are dense, their underlying rank structure is mathematically equivalent to being almost-banded and the systems of equations can still be solved in linear time.

2. A simple example. We present the main idea by means of a particular example, namely, given a continuous function f defined on $[-1, 1]$, find u such that

$$u' + u = f, \quad u(1) = 0.$$

As is well established in [13], this model serves as a prototype for a larger class of problems. Our first task is to rewrite this equation in terms of coefficients of orthogonal polynomials. We will do this in two ways. The first method will construct the matrices using the techniques in [13]. The second will construct equivalent matrices that depend only on Chebyshev coefficients and involve no conversions to other ultraspherical representations. When clarity is needed, we subscript variables corresponding to the first method by "ultra" and variables corresponding to the second method with "cheb".

Our linear operator can be written as the sum of two basic operators: differentiation (D) and identity (I). In the first method, the ultraspherical method of Olver and Townsend, D is the matrix that maps the Chebyshev coefficients of u to the ultraspherical coefficients of u'.

*Received April 21, 2015. Accepted May 8, 2015. Published online on May 18, 2015. Recommended by L. Reichel. This research was partially supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 291068. The views expressed in this article are not those of the ERC or the European Commission, and the European Union is not liable for any use that may be made of the information contained here.

†Mathematical Institute, University of Oxford, Andrew Wiles Building, Woodstock Road, OX2 6GG Oxford, UK (aurentz@maths.ox.ac.uk).

[1] This is true when the coefficients of the differential equation are low-degree polynomials.

[2] Almost-banded matrices also appear in the Chebyshev-based spectral methods in [3, 9]. These techniques rely on first reformulating the differential equation as an integral equation.

ETNA
Kent State University
http://etna.math.kent.edu

Let $\{u_i\}_{i=0}^{\infty}$ be the Chebyshev coefficients of u. The matrix D_{ultra} is defined as

$$D_{\text{ultra}} u = \begin{bmatrix} 0 & 1 & & & \\ & & 2 & & \\ & & & 3 & \\ & & & & \ddots \\ & & & & \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ \vdots \end{bmatrix}.$$

Similarly, one can write down the matrix I_{ultra} that takes Chebyshev coefficients to ultraspherical coefficients (also known as the conversion operator),

$$I_{\text{ultra}} u = \frac{1}{2} \begin{bmatrix} 2 & 0 & -1 & & \\ & 1 & 0 & -1 & \\ & & 1 & 0 & \\ & & & & \ddots \\ & & & & \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ \vdots \end{bmatrix}.$$

To account for the boundary condition, we need an additional functional for computing $u(1)$. For Chebyshev coefficients this takes the form of a simple inner product,

$$u(1) = \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ \vdots \end{bmatrix}.$$

Combining all these pieces together, we get the following system of equations:

$$\frac{1}{2} \begin{bmatrix} 2 & 2 & 2 & 2 & \cdots \\ 2 & 2 & -1 & & \\ & 1 & 4 & -1 & \\ & & 1 & 6 & \\ & & & & \ddots \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{f}_0 \\ \hat{f}_1 \\ \hat{f}_2 \\ \vdots \end{bmatrix}. \tag{2.1}$$

Here the values $\{\hat{f}_i\}$ are the ultraspherical coefficients of f computed by multiplying f by I_{ultra}. We will refer to this matrix as L_{ultra}.

To facilitate our discussion, we focus on the nonzero pattern. Below is the almost-banded pattern of the 10×10 L_{ultra} matrix,

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & & & & & & & \\ & \times & \times & \times & & & & & & \\ & & \times & \times & \times & & & & & \\ & & & \times & \times & \times & & & & \\ & & & & \times & \times & \times & & & \\ & & & & & \times & \times & \times & & \\ & & & & & & \times & \times & \times & \\ & & & & & & & \times & \times & \times \\ & & & & & & & & \times & \times \end{bmatrix}. \tag{2.2}$$

ETNA
Kent State University
http://etna.math.kent.edu

Before we start discussing equivalence, we need the matrices for the second method, the more standard spectral method based on Chebyshev coefficients; see, e.g., [11]. The matrices D and I now take the following forms:

$$D_{\text{cheb}} = \begin{bmatrix} 0 & 1 & & 3 & & 5 & & 7 & & \cdots \\ & & 4 & & 8 & & 12 & & 16 & \\ & & & 6 & & 10 & & 14 & & \\ & & & & 8 & & 12 & & 16 & \\ & & & & & \ddots & & & & \ddots \end{bmatrix}, \qquad I_{\text{cheb}} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & \ddots \end{bmatrix}.$$

Notice how all the rows of D_{cheb} with even/odd indices are multiples of each other away from the main diagonal. This hints at a special structure that can be used to describe D_{cheb} using only $O(n)$ parameters.

The entire system expressed as a single matrix has the following form:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \cdots \\ 1 & 1 & & 3 & & 5 & & 7 & & \cdots \\ & 1 & 4 & & 8 & & 12 & & 16 & \\ & & 1 & 6 & & 10 & & 14 & & \\ & & & 1 & 8 & & 12 & & 16 & \\ & & & & \ddots & & & & & \ddots \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 \\ f_0 \\ f_1 \\ f_2 \\ f_3 \\ \vdots \end{bmatrix}. \tag{2.3}$$

We will refer to this matrix as L_{cheb}.

Here we can see why using only Chebyshev coefficients looks expensive. The new system is essentially dense in the upper-triangular part and appears to require $O(n^2)$ storage. Below we illustrate the nonzero pattern of the 10×10 L_{cheb} matrix,

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\ \times & \times & & \times & & \times & & \times & & \times \\ & \times & \times & & \times & & \times & & \times & \\ & & \times & \times & & \times & & \times & & \times \\ & & & \times & \times & & \times & & \times & \\ & & & & \times & \times & & \times & & \times \\ & & & & & \times & \times & & \times & \\ & & & & & & \times & \times & & \times \\ & & & & & & & \times & \times & \\ & & & & & & & & \times & \times \end{bmatrix}. \tag{2.4}$$

3. Rank-structured matrices. Now that we have our matrices, (2.1)–(2.2) and (2.3)–(2.4) in hand, we can describe how these two have essentially the same *rank structure*. Rank-structured matrices are quite prevalent in applied mathematics, and in the last few decades considerable work has been done to develop efficient algorithms. The central idea is that even dense matrices can be represented by very few parameters if certain submatrices have small rank. Let us look more closely at L_{ultra}. Below we have highlighted a rectangular submatrix whose lower-left corner intersects the main diagonal,

$$\begin{bmatrix}
\times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & & & & & & & \\
& \times & \times & \times & & & & & & \\
& & \times & \times & \times & & & & & \\
& & & \times & \times & \times & & & & \\
& & & & \times & \times & \times & & & \\
& & & & & \times & \times & \times & & \\
& & & & & & \times & \times & \times & \\
& & & & & & & \times & \times & \times \\
& & & & & & & & \times & \times
\end{bmatrix}.$$

Notice that the highlighted matrix has rank at most 3. This remains true for any submatrix whose lower-left corner intersects the main diagonal,

$$\begin{bmatrix}
\times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & & & & & & & \\
& \times & \times & \times & & & & & & \\
& & \times & \times & \times & & & & & \\
& & & \times & \times & \times & & & & \\
& & & & \times & \times & \times & & & \\
& & & & & \times & \times & \times & & \\
& & & & & & \times & \times & \times & \\
& & & & & & & \times & \times & \times \\
& & & & & & & & \times & \times
\end{bmatrix},$$

$$\begin{bmatrix}
\times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & \times & & & & & & & \\
& \times & \times & \times & & & & & & \\
& & \times & \times & \times & & & & & \\
& & & \times & \times & \times & & & & \\
& & & & \times & \times & \times & & & \\
& & & & & \times & \times & \times & & \\
& & & & & & \times & \times & \times & \\
& & & & & & & \times & \times & \times \\
& & & & & & & & \times & \times
\end{bmatrix}.$$

Now we draw a similar box on the L_{cheb} matrix,

$$\begin{bmatrix}
\times & \times & \times & \times & \times & \times & \times & \times & \times & \times \\
\times & \times & & \times & & \times & & \times & & \times \\
& \times & \times & & \times & & \times & & \times & \\
& & \times & \times & & \times & & \times & & \times \\
& & & \times & \times & & \times & & \times & \\
& & & & \times & \times & & \times & & \times \\
& & & & & \times & \times & & \times & \\
& & & & & & \times & \times & & \times \\
& & & & & & & \times & \times & \\
& & & & & & & & \times & \times \\
& & & & & & & & & \times
\end{bmatrix}.$$

This submatrix too has rank at most 3! In fact, any submatrix whose lower-left corner intersects the main diagonal will have rank at most 3. To see this, we decompose L_{cheb} into three parts and highlight the same submatrix,

$$L_{\text{cheb}} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & & & & & & & & \\ & 1 & 0 & & & & & & & \\ & & 1 & 0 & & & & & & \\ & & & 1 & 0 & & & & & \\ & & & & 1 & 0 & & & & \\ & & & & & 1 & 0 & & & \\ & & & & & & 1 & 0 & & \\ & & & & & & & 1 & 0 & \\ & & & & & & & & 1 & 0 \end{bmatrix} +$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 1 & & 3 & & 5 & & 7 & & 9 \\ & & 0 & & & & & & & \\ & & & 6 & & 10 & & 14 & & 18 \\ & & & & 0 & & & & & \\ & & & & & 10 & & 14 & & 18 \\ & & & & & & 0 & & & \\ & & & & & & & 14 & & 18 \\ & & & & & & & & 0 & \\ & & & & & & & & & 18 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ & 0 & & & & & & & & \\ & & 4 & & 8 & & 12 & & 16 & \\ & & & 0 & & & & & & \\ & & & & 8 & & 12 & & 16 & \\ & & & & & 0 & & & & \\ & & & & & & 12 & & 16 & \\ & & & & & & & 0 & & \\ & & & & & & & & 16 & \\ & & & & & & & & & 0 \end{bmatrix}.$$

Since the highlighted submatrix in each of the 3 parts has rank at most 1, the highlighted block of L_{cheb} has rank at most 3.

This type of rank structure is a generalization of *semiseparability*.[3] The study of rank-structured matrices originates with Gantmacher and Kreĭn [10]. For more recent developments, see, for example, [4, 15].

4. Fast solutions via QR factorization. In [13] the discretized systems are solved using QR factorizations. The authors use the almost-bandedness to show that both of the factors Q and R in $QR = L_{\text{ultra}}$ have a special structure that makes it possible to compute Q and perform the back solve in linear time. We now observe that it is not the almost-bandedness that matters, but it is really the rank structure. This means that it is just as easy to develop a linear time QR-based solver for the L_{cheb} system.

The details are a bit complicated and will be spelled out with numerical applications in a future publication. Here we highlight the key points. The first thing to note is that both L_{ultra} and L_{cheb} are upper-Hessenberg. This implies that Q_{ultra} and Q_{cheb} ($QR = L$) are upper-Hessenberg. Consequently, they and their inverses can be represented by products of n elementary 2×2 matrices and applied to a vector in linear time.

[3] Applicable definitions include both *sequentially-semiseparable* [5] and *quasiseparable* [7].

The R matrices are much more interesting. In general they are dense,

$$Q_{\text{ultra}}^{*} L_{\text{ultra}} = R_{\text{ultra}} = \begin{bmatrix} \times & \times & \times & \times & \times & \times & \times & \times & \times & \cdots \\ & \times & \times & \times & \times & \times & \times & \times & \times & \\ & & \times & \times & \times & \times & \times & \times & \times & \\ & & & \times & \times & \times & \times & \times & \times & \\ & & & & \times & \times & \times & \times & \times & \\ & & & & & \times & \times & \times & \times & \\ & & & & & & \times & \times & \times & \\ & & & & & & & \ddots & & \end{bmatrix},$$

but any submatrix whose lower-left corner intersects the diagonal has rank at most 4,

$$\begin{bmatrix} \times & \times & \times & \times & \times & \times & \times & \times & \times & \cdots \\ & \times & \times & \times & \times & \times & \times & \times & \times & \\ & & \times & \times & \times & \times & \times & \times & \times & \\ & & & \times & \times & \times & \times & \times & \times & \\ & & & & \times & \times & \times & \times & \times & \\ & & & & & \times & \times & \times & \times & \\ & & & & & & \times & \times & \times & \\ & & & & & & & \ddots & & \end{bmatrix}.$$

One can explain the rank structure by recalling that L_{ultra} is rank-structured and unitary matrices preserve rank. Exactly the same reasoning can be applied to L_{cheb} with the same conclusion.

The back substitution, in both methods, involves an upper-triangular matrix with a definite rank structure. Olver and Townsend exploited this structure to develop a linear time algorithm for performing a back solve with R_{ultra}. Their algorithm is an extension of a similar method that was first developed by Chandrasekaran and Gu [2]. The same technique can also be used for R_{cheb}.

4.1. Conditioning. It should be noted that the simple form of the Chebyshev differentiation matrices presented here is subject to the same ill-conditioning as other Chebyshev-based spectral methods [1, 8]. In [13] the authors construct a structure-preserving preconditioner to ensure that the ultraspherical-based systems are well conditioned. A similar technique can be applied to the Chebyshev-based methods proposed in this note.

5. Other orthogonal families. Above, we considered a specific Chebyshev example. We conclude this note by showing that orthogonal families more generally possess rank-structured differentiation matrices. The matrices for these examples are derived from formulas in [14].

The first example is the 10×10 version of L_{leg}, the spectral discretization of $u' + u = f$,

ETNA
Kent State University
http://etna.math.kent.edu

with $u(1) = 0$, using Legendre coefficients,

$$L_{\text{leg}} = \left[\begin{array}{ccccc|ccccc|}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & & 1 & & 1 & & 1 & & 1 \\
& 1 & 3 & & 3 & & 3 & & 3 & \\
& & 1 & 5 & & 5 & & 5 & & 5 \\
& & & 1 & 7 & & 7 & & 7 & \\
& & & & 1 & 9 & & 9 & & 9 \\ \cline{6-10}
& & & & & 1 & 11 & & 11 & \\
& & & & & & 1 & 13 & & 13 \\
& & & & & & & 1 & 15 & \\
& & & & & & & & 1 & 17
\end{array}\right].$$

Just as with Chebyshev bases, this submatrix has rank at most 3. A similar structure also exists for trigonometric polynomials as illustrated by our second example. Below is the 9×9 version of L_{four}, the spectral discretization of $u' = f,\ u(0) = 0$ with $\int f = 0$, using Fourier coefficients,

$$L_{\text{four}} = \left[\begin{array}{cccc|ccccc|}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
& 1i & & & & & & & \\
& & -1i & & & & & & \\
& & & 2i & & & & & \\
& & & & -2i & & & & \\ \cline{5-9}
& & & & & 3i & & & \\
& & & & & & -3i & & \\
& & & & & & & 4i & \\
& & & & & & & & -4i
\end{array}\right].$$

Here the rank is at most 2. The low-rank structure can even be observed in spectral methods for unbounded domains. Our last example is the 10×10 version of L_{lag}, the spectral discretization of $u' + u = f$, $u(0) = 0$, using Laguerre coefficients,

$$L_{\text{lag}} = \left[\begin{array}{ccccc|ccccc|}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
& 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
& & 1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \\
& & & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\
& & & & 1 & -1 & -1 & -1 & -1 & -1 \\ \cline{6-10}
& & & & & 1 & -1 & -1 & -1 & -1 \\
& & & & & & 1 & -1 & -1 & -1 \\
& & & & & & & 1 & -1 & -1 \\
& & & & & & & & 1 & -1
\end{array}\right].$$

Here the rank is again bounded by 2.

Conclusions. It has been shown that spectral differentiation matrices possess structure that can be exploited to develop fast methods for a variety of orthogonal families. This observation opens the door to new classes of spectral methods for solving ordinary and partial differential equations.

Acknowledgements. The author would like to thank the editor and referee for their comments which helped improve the quality of this note and for their incredibly fast response time. The author would also like to thank Nick Trefethen for his help with both content and aesthetics and Raf Vandebril for his help clarifying various technical details.

REFERENCES

[1] J. P. BOYD, *Chebyshev and Fourier Spectral Methods*, Dover, Mineola, 2001.
[2] S. CHANDRASEKARAN AND M. GU, *Fast and stable algorithms for banded plus semiseparable systems of linear equations*, SIAM J. Matrix Anal. Appl., 25 (2003), pp. 373–384.
[3] C. W. CLENSHAW, *The numerical solution of linear differential equations in Chebyshev series*, Proc. Cambridge Philos. Soc., 53 (1957), pp. 134–149.
[4] S. DELVAUX, *Rank Structured Matrices*, Ph.D. Thesis, Dept. of Computer Science, KU Leuven, Leuven, 2007.
[5] P. DEWILDE AND A.-J. VAN DER VEEN, *Time-Varying Systems and Computations*, Springer, Dordrecht, 1998.
[6] T. A. DRISCOLL, N. HALE, AND L. N. TREFETHEN, *Chebfun Guide*, Pafnuty Publications, Oxford, 2014.
[7] Y. EIDELMAN AND I. GOHBERG, *On a new class of structured matrices*, Integral Equations Operator Theory, 34 (1999), pp. 293–324.
[8] B. FORNBERG, *A Practical Guide to Pseudospectral Methods*, Cambridge University Press, Cambridge, 1996.
[9] L. FOX AND I. B. PARKER, *Chebyshev Polynomials in Numerical Analysis*, Oxford University Press, London, 1968.
[10] F. R. GANTMACHER AND M. G. KREĬN, *Sur les matrices oscillatoires et complètement non négatives*, Compos. Math., 4 (1937), pp. 445–476.
[11] A. KARAGEORGHIS, *A note on the Chebyshev coefficients of the general order derivative of an infinitely differentiable function*, J. Comput. Appl. Math., 21 (1988), pp. 129–132.
[12] S. OLVER, G. GORETKIN, R. M. SLEVINSKY, AND A. TOWNSEND, *ApproxFun.jl*, 2015. https://github.com/ApproxFun/ApproxFun.jl
[13] S. OLVER AND A. TOWNSEND, *A fast and well-conditioned spectral method*, SIAM Rev., 55 (2013), pp. 462–489.
[14] J. SHEN, T. TANG, AND L.-L. WANG, *Spectral Methods: Algorithms, Analysis and Applications*, Springer, Heidelberg, 2011.
[15] R. VANDEBRIL, M. VAN BAREL, AND N. MASTRONARDI, *Matrix Computations and Semiseparable Matrices, Vol. 1*, Johns Hopkins University Press, Baltimore, 2007.