[Review published in *SIAM Review*, Vol. 58, Issue 1, pp. 159–163.]

Featured Review: Approximation Theory and Approximation Practice. *By Lloyd N. Trefethen.* SIAM, Philadelphia, 2013. $51.00. viii+305 pp., softcover. ISBN 978-1-611972-39-9.

Approximation Theory and Approximation Practice (hereinafter ATAP) is another remarkable book from Nick Trefethen. ATAP begins with "Welcome to a beautiful subject!—the constructive approximation of functions. And welcome to a rather unusual book." To which this reviewer replies, "And how."

The first unusual thing is that the book was produced using MATLAB's "publish" feature, and each chapter is available for download free of charge, complete with all source code from `www.chebfun.org/ATAP/` . This represents an exemplary commitment to reproducible research and to open access [11]. SIAM publishing deserves an accolade for going along with this. Of course there's the omnipresence of piracy to think of. How better to disarm the pirates than by giving away what they want to steal? Now, SIAM is so active in supporting our research and teaching that I, for one, had no hesitation in actually purchasing a copy anyway—and I'm certain I'm not alone.

Other publishers are trying to cope too. Springer provides free electronic access to members of institutions that have purchased the Springer Electronic Library and moreover sells universally priced copies ($25) of each book in that list. The CUP book *Modern Computer Arithmetic* by Brent and Zimmerman [1] can be downloaded for free—well, the last prepublication draft anyway—at `http://arxiv.org/pdf/1004.4710.pdf`. So there are other models. But I like SIAM's model. It's clean, and is not just open access but also, as I said, reproducible research.

Next, and this is a big change from how ordinary people write books, Nick cites the original sources (with an annotated bibliography, no less). Nick is a self-described history buff. One might even say *scholar*. Most of us are lousy at history (well, I am anyway). If someone calls something the Gibbs phenomenon, I just believe them, and never find out that it was Wilbraham who did it first.[1] Nick's historical care adds a graceful tone to ATAP, and his authoritative sleuthing—though he probably enjoyed it—is of immense value. The only comparably accurate books I know are Hairer and Wanner's *Analysis by Its History* [7] and Ostermann and Wanner's *Geometry by Its History* [10]. If the only thing ATAP did was just set the record straight, it would still be worth buying for that alone.

But it does a great deal more than that. For one thing, most of its theorems come instantiated or illustrated in Chebfun, which is a MATLAB package that can be downloaded again free of charge from `www.chebfun.org`. Chebfun is a project of some maturity now and was begun with the goal of bringing the speed of floating-point computation to the problem of computing with functions. It succeeds remarkably well. The success of this project comes mostly from its basic design decision, to represent univariate real functions by Chebyshev series, or equivalently numerically by interpolation at Chebyshev–Lobatto aka Chebyshev extreme points $x_k = \cos(\pi k/n)$ on the interval $[-1, 1]$. The approximation theory described in ATAP thus plays a fundamental role in the design of Chebfun, and so it's natural that Chebfun be used to illustrate approximation practice. A great many impressive examples of Chebfun in action can be found at `www.chebfun.org/examples/`, by the way.

An example of my own that illustrates the utility of Chebfun to resurrect a very old method for solving delay differential equations, namely, the method of steps, can be found in Chapter 16 of [2]. See Figure **??**. The method of steps fails in symbolic computation, usually when the integrals become nonelementary; but even if the integrals are all just polynomials, the complexity suffers greatly from the swelling size of the coefficients and the length of the series. In Chebfun, essentially all functions are repre-

[1] Well, now I know this, from reading ATAP.

sented by Chebyshev series truncated automatically at roundoff-level accuracy; such series can be integrated repeatedly without either significant error or swell in size. The degrees of the approximation might be in the thousands, or even tens of thousands, but this presents no difficulty.

This brings up a myth, one of several that ATAP explicitly overturns: the myth that polynomial interpolants diverge as the degree goes to infinity. It's a pretty powerful myth, because it has a theorem [4], which says no matter what the grids are, there is a function f for which the interpolants diverge. But if f is even just Lipschitz continuous, convergence is guaranteed if the Chebyshev nodes are used, and the smoother f is, the faster. This matters enormously, in practice. These facts are discussed in detail in Chapters 7 and 8 of ATAP. These facts also have important implications for numerical quadrature (Chapter 19).

Of course the use of high-degree polynomial interpolants requires a numerically stable algorithm for evaluating them, namely, the first and second barycentric forms, which are much better than divided differences (another myth overturned!). There is another reason divided differences are not as good, and I'll return to this.

There are 28 chapters in ATAP. My favorites are Chapter 26: "Rational Interpolation and Least Squares" (which has a strong connection to my work in symbolic-numeric algorithms for polynomials, or at least it will), Chapter 12: "Potential Theory and Approximation," and Chapter 23: "Nonlinear Approximation: Why Rational Functions?"

That last chapter is perhaps the keystone of the book. It explains, really explains, why approximation by polynomials or by rationals is interesting or useful at all: not for the often-stated reasons to do with evaluation of functions on computers, but because $r(A)$ is easier to evaluate than $\exp(A)$, where A is a matrix or operator; not because we can evaluate $p(x)$, but because we can find its roots.

In some sense that's a spoiler. I've nearly quoted Nick verbatim there. But there's a reason this is Chapter 23 and not Chapter 1: those statements don't carry their full weight without the 22 prior chapters!

Chapter 12 is one of the most beautiful. It makes a connection between potential theory and approximation theory that gives a very satisfying explanation for what differentiates a good set of nodes for approximation from a poor set. There's a lot to say, and the chapter is quite concise and bears rereading, as Nick himself notes on p. 91.

Now I'm going to make a connection between ATAP and Nick's work on pseudospectra with the point of view taken in some of my own work (including my paper [3] that Nick cites in ATAP, which has been taken up in more detail in section 8.4 of [2]). This also further justifies the omission of divided differences. The point is that ATAP is written with the idea in mind that there is no knowing which polynomial will wind up being used, so one ought to be concerned with Lebesgue constants (Chapter 15 of ATAP) and not condition numbers of individual polynomials. In contrast, my work follows the ideas of [5], [6], and [9] and is concerned with condition numbers. This work is very similar in spirit to the work on pseudospectra by Nick Trefethen and coworkers, and indeed there's an explicit connection.

Here's what I mean. We follow the definitions of [6]. If

$$p(z) = \sum_{k=0}^{n} c_k \phi_k(z)$$

is an expression for a polynomial (or matrix polynomial if the c_k are matrices) in some basis ϕ, then making small relative changes in the coefficients c_k, say to $c_k(1+\delta_k)$, results in a change in the value of $p(z)$ to $p(z)$ + $\Delta p(z)$ where

$$\Delta p(z) = \sum_{k=0}^{n} c_k \delta_k \phi_k(z)\,.$$

This is at most (by the triangle inequality)

$$|\Delta p(z)| \le \left(\sum_{k=0}^{n} |c_k||\phi_k(z)| \right) \cdot \|\delta\|_\infty,$$

and the (nonpolynomial) function

$$B(z) = \sum_{k=0}^{n} |c_k||\phi_k(z)|$$

acts as a condition number. The bound is attained for maximal, signed δ_k. This condition number of p depends on the basis ϕ_k

in two ways—one through ϕ_k itself, and another in that the coefficients c_k themselves change if you change the basis. Farouki and Goodman [5] give an elegant proof that Bernstein bases

$$\phi_j(z) = (b-a)^{-n}\binom{n}{j}(x-a)^j(b-x)^{n-j}$$

are, among all nonnegative bases on an interval, optimal in the sense that no expression in terms of other bases nonnegative on the interval can have generally better condition number. The paper [3] weakens the nonnegativity constraint and shows that Lagrange bases can be better yet, often by many orders of magnitude.

This is not true for Newton bases (divided differences). Additionally their condition number depends unhelpfully on the ordering of the nodes. The point of [3] is, however, that a polynomial expressed in a Lagrange basis can be expected to be much better conditioned than the same polynomial expressed in a Newton basis, that is, by use of divided differences. This to me is the real reason Lagrange bases are better than divided differences. This also helps Chebfun when the Lagrange basis on Chebyshev-nodes is used.

ATAP doesn't use the full strength of this argument, because it (the argument) depends on the particular polynomial. Instead, ATAP relies on the more generic, but weaker, bound that arises from

$$B(z) \le \left(\sum_{k=0}^{n} |\phi_k(z)|\right)\|c\|_\infty,$$

where the Lebesgue function

$$\Lambda(z) = \sum_{k=0}^{n} |\phi_k(z)|$$

and the Lebesgue constant $\Lambda = \sup\big(\Lambda(z)\big)$ give larger and larger bounds. This weakening can be quite consequential, although it does make general a priori conclusions possible. More importantly, because the Lebesgue constant for Chebyshev bases is nearly optimal, the Chebyshev–Lobatto interpolation nodes $y_k = \cos(\pi k/n)$, $0 \le k \le n$, are also extremely good.

Interestingly, the function $B(z)$ arises in the computation of pseudospectra or pseudozeros: using the symbol Λ differently,

$$\Lambda_\varepsilon(z) := \left\{ z \;\middle|\; \exists \Delta c_k \text{ with } |\Delta c_k| \le \alpha_k \varepsilon \text{ and } \sum_{k=0}^{n}(c_k + \Delta c_k)\phi_k(z) = 0 \right\}$$

is the pseudozero set for $p(z)$, with nonnegative weights α_k (typically $\alpha_k = |c_k|$ for relative weights). An alternative characterization for Λ_ε is familiar:

$$\Lambda_\varepsilon(z) := \{z \mid |p(z)| \le B(z)\varepsilon\}$$

with our previously defined $B(z)$ (in the matrix case one has $\|P^{-1}(z)\| \ge (B(z)\varepsilon)^{-1}$). For pseudospectra in general of course the go-to reference is [12].

I think this is an important part of approximation theory that is not in ATAP. On the other hand, Lebesgue constants are important, too, and they can be used a priori, whereas condition numbers can only be used a posteriori. I mention Lebesgue constants only once in [2], and in retrospect that's not enough. So maybe Nick's right and I'm wrong. Speaking of errors, there are inevitably some in ATAP. But known errors are listed on the well-maintained errata page at `www.chebfun.org/ATAP/` and if you find another, just send it to Nick.

Much more important are the exercises. They are novel, interesting, challenging, informative and in some cases open-ended. For instance, you might find yourself learning a lot about order stars [8] if you really put your mind to answering Exercise 27.4. I confess that I have not done all the exercises. I did use this review as an excuse to do some more. They're fun as well as challenging. Don Knuth once said (either to me or to David Jeffrey, I forget which) that he included the answers to all the exercises in his books because he worried that he "might not be able to do them again." I don't generally approve of having the answers to past exams available for students—but I would read Nick's answers to his exercises, should he ever publish them.

I was warned that reviews have a tendency to turn into essays, and I guess this happened and I've strayed a bit from my topic. I'll just finish by stating what should be obvious: ATAP is a landmark, revolutionary book.

Acknowledgments. This review was prepared with the help of Torin Viger

and Julia E. Jankowski, who updated the code to generate Figure **??** and helped with the preparation of the manuscript. I also thank David Watkins for the opportunity to write this review, and for his helpful comments on an earlier draft.

REFERENCES

[1] R. Brent and P. Zimmermann, *Modern Computer Arithmetic*, Cambridge Monogr. Appl. Comput. Math. 18, Cambridge University Press, Cambridge, UK, 2011.

[2] R. M. Corless and N. Fillion, *A Graduate Introduction to Numerical Methods: From the Viewpoint of Backward Error Analysis*, Springer-Verlag, New York, 2013.

[3] R. M. Corless and S. M. Watt, *Bernstein bases are optimal, but, sometimes, Lagrange bases are better*, in Proceedings of the 6th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), Timisoara, 2004, pp. 141–153.

[4] G. Faber, *Über die interpolatorische Darstellung stetiger Funktionen*, Jahresbericht der Deutschen Mathematiker-Vereinigung, 23 (1914), pp. 190–210.

[5] R. T. Farouki and T. N. T. Goodman, *On the optimal stability of the Bernstein basis*, Math. Comp., 65 (1996), pp. 1553–1566.

[6] R. T. Farouki and V. Rajan, *On the numerical condition of polynomials of Bernstein form*, Comput. Aided Geom. Design, 4 (1987), pp. 191–216.

[7] E. Hairer and G. Wanner, *Analysis by Its History*, Springer-Verlag, New York, 1996.

[8] A. Iserles and S. Nørsett, *Order stars and rational approximants to* $\exp(z)$, Appl. Numer. Math., 5 (1989), pp. 63–70.

[9] T. Lyche and J. M. Peña, *Optimally stable multivariate bases*, Adv. Comput. Math., 20 (2004), pp. 149–159.

[10] A. Ostermann and G. Wanner, *Geometry by Its History*, Springer-Verlag, Heidelberg, 2012.

[11] V. Stodden and S. Miguez, *Best Practices for Computational Science: Software Infrastructure and Environments for Reproducible and Extensible Research*, article 2322276, Social Science Research Network, 2013; available online from `http://papers.ssrn.com/sol3/papers.cfm?abstract`$_i d = 2322276$.

[12] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, Princeton, NJ, 2005.

Robert M. Corless
University of Western Ontario