From Self-Consistency to SOAR:
Solving Large-Scale Nonlinear Eigenvalue Problems

By Zhaojun Bai and Chao Yang

What do electronic structure calculations, the design of MEMS devices, vibrational analysis of high-speed railways, and simulation of the electromagnetic field of a particle accelerator have in common? Answer: All require the solution of large-scale nonlinear eigenvalue problems. In fact, these are just a handful of areas in which accurate, efficient solutions of nonlinear eigenvalue problems are playing an increasingly prominent role.

Recognizing the importance of this class of problems, the organizers of the 2005 SIAM Annual Meeting included an invited minisymposium on nonlinear eigenvalue problems in the program. The two-part (eight-talk) minisymposium brought together numerical analysts and application scientists; together, the sessions featured cutting-edge results from both communities, along with discussions of ongoing challenges.

Electronic Structure Calculations

The first three speakers discussed a nonlinear eigenvalue problem that arises in electronic structure calculations. In problems of this type, the matrix Hamiltonian H depends, in a nontrivial way, on the set of eigenvectors X to be computed. The invariant subspace spanned by these eigenvectors minimizes a total energy function that is highly nonlinear with respect to X on a manifold defined by a set of orthonormality constraints.

The most widely used method for tackling the electronic structure problem is the *self-consistent field* (SCF) iteration. The method solves a sequence of linear eigenvalue problems associated with the fixed Hamiltonian $H(X_i)$, where X_i is the approximation to X obtained in the previous iteration. At convergence, the difference be-tween $H(X_i)$ and $H(X_{i-1})$ is negligible, i.e., the eigenvectors associated with the desired eigenvalues of $H(X_i)$ have become *self-consistent*. Convergence is not guaranteed, however, and, in fact, for many systems the SCF iteration fails to converge.

Chao Yang, a staff scientist at Lawrence Berkeley National Laboratory, presented an alternative approach that he has developed with LBNL colleagues Juan Meza and Ling-wang Wang. Their approach obtains the desired eigenvectors (see Figure 1) through a direct minimization of the total energy function. The algorithm they use projects the total energy function into a sequence of subspaces of small dimensions. By minimizing the total energy function within each of these subspaces, they are able to avoid solving a sequence of large linear eigenvalue problems. Minimizing the total energy function directly also prevents some of the convergence problems encountered in SCF.

Robert Harrison, a computational chemist at Oak Ridge National Laboratory and the University of Tennessee, Knoxville, and one of the original developers of the widely used chemistry package NWChem, presented another novel approach to electronic structure calculation. The approach, developed in collaboration with Gregory Beylkin of the University of Colorado, Boulder, and George Fann of ORNL, reformulates the problem as an integral equation and avoids many of the difficult numerical issues associated with the differential operator. In particular, the smallest eigenvalues of the differential operator become the largest eigenvalues of the integral operator (which still depends nonlinearly on the eigenvectors to be computed). These eigenvalues tend to converge more rapidly. Furthermore, because of the smooth and rapid decay of the Green's function associated with the bound-state Helmholtz equation used in this formulation, the integral kernel can be represented by a multi-wavelet basis. As a result, the convolution computation can be carried out efficiently with special multi-resolution techniques.

Figure 1. *Isosurface rendering of an eigenvector associated with an eigenvalue in the conducting band of a silicon quantum dot. The larger and darker balls in the figure correspond to silicon atoms. The smaller and lighter balls correspond to hydrogen atoms. Image courtesy of Lin-wang Wang of LBNL.*

For certain types of nanoscale structures, such as semiconductor pyramid quantum dots in which the carriers are confined in all three dimensions, the electronic structure calculation can be simplified to a rational eigenvalue problem in which the matrix Hamiltonian is a rational function of the eigenvalues to be computed. Heinrich Voss of the Hamburg University of Technology, in Germany, discussed new methods for solving such problems. In particular, he presented a nonlinear Arnoldi algorithm, the Jacobi–Davidson (JD) algorithm, and a modified version of the JD algorithm. He compared the performance of these approaches, and pointed out the need to take advantage of the minimax characterization of the nonlinear eigenvalue problem when symmetry is present in the problem. Elaborating on Voss's discussion of the JD algorithm in the solution of general nonlinear eigenvalue problems, Henk Van der Vorst of Utrecht University, in the Netherlands, provided theoretical background as well as numerical examples in which the JD algorithm was used to solve polynomial eigenvalue problems.

MEMS Devices

David Bindel, a graduate student at the University of California, Berkeley, presented a new approach for solving a complex symmetric *quadratic* eigenvalue problem arising in the simulation of resonance in microelectronic mechanical systems (MEMS). High-frequency electromechanical resonators are important components of modern communication systems, such as radiofrequency ID tags and cell phones. With recent progress in surface micromachining technology, it is possible that, in the near future, we will have "Dick Tracy" phones that are smaller and require less power than today's cell phones.
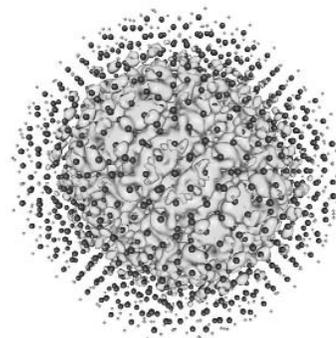
For these designs to work, it is necessary to minimize the amount of damping in the resonating components. While engineers have several tools for estimating the resonant frequencies of these devices, they have far fewer tools for estimating damping. Calculations of the damped resonances of such coupled-domain devices typically involve nonlinear eigenvalue problems with a variety of special structures. (See Figure 2.)

Figure 2. *Cutaway schematic (left) and SEM picture (right) of a micromachined disk resonator. Through modulated electrostatic attraction between a disk and a surrounding ring of electrodes, the disk is driven into mechanical resonance. Because the disk is anchored to a silicon wafer, energy leaks from the disk to the substrate, from which it radiates away as elastic waves. To study this energy loss, David Bindel has constructed finite element models in which the substrate is modeled by a perfectly matched absorbing layer. Resonance poles are approximated by eigenvalues of a large, sparse complex-symmetric matrix pencil. For more details, see D.S. Bindel and S. Govindjee, "Anchor Loss Simulation in Resonators,"* International Journal for Numerical Methods in Engineering, *Vol. 64, No. 6, 2005. Resonator micrograph courtesy of Emmanuel Quevy.*

Preserving Special Matrix Structures

In the second session, Christian Mehl of Technical University of Berlin, Germany, described numerical techniques for solving a special type of polynomial eigenvalue problem arising in the vibration analysis of rail tracks excited by high-speed trains. This problem—also highlighted in the invited talk given at the meeting by Volker Mehrmann, also of TU Berlin—was described in the article "Accurate Eigenvalues for Fast Trains" by Ilse Ipsen in the November 2004 issue of *SIAM News*.

Yangfeng Su of Fudan University, China, presented a structure-preserving method for computing eigenvalues of skew-Hamilton-ian/Hamiltonian (SHH) pencils. Eigenvalue problems of this type arise in the passivity analysis of VLSI circuits and from complex gyroscopic quadratic eigenvalue problems. Lie-quan Lee, a computational scientist at the Stanford Linear Accelerator Center, discussed nonlinear eigenvalue problems encountered by designers of accelerator cavities; this project is funded by the U.S. Department of Energy's Scientific Discovery through Advanced Computing program.

These three talks had a common theme: Many nonlinear eigenvalue problems possess special matrix structures and spectral properties. The polynomial eigenvalue problems that come up in the rail track vibrational analysis, for example, are known as *palindromic* eigenvalue problems because of the symmetric pattern of the polynomial. As a result, eigenvalues appear in a reciprocal pair, i.e., if λ is an eigenvalue, so is λ^{-1}.

Although some of these problems can be reduced or transformed into equivalent linear eigenvalue problems, applying the standard algorithm to the linear problem often destroys the special properties of the eigenvalues associated with the original nonlinear problem. There is thus a need for numerical algorithms that can preserve the structures of the original problems and desired eigenvalues. Mehl discussed a combination of a Jacobi-like algorithm and a technique developed for computing the Hamiltonian–Schur form to solve palindromic eigenvalue problems. Su presented a novel method that projects the eigenproblem of SHH pencils onto a special pair of subspaces in order to preserve the SHH pencil structure.

Lee described the use of SLAC's Omega3P package for simulating the design of an accelerator cavity with external coupling. The package solves a quadratic eigenvalue problem by a second-order Arnoldi (SOAR) method, recently proposed by Zhaojun Bai and Yangfeng Su (*SIAM Journal on Matrix Analysis and Applications*, Vol. 26, No. 3, 2005), and the fully nonlinear problem by self-consistent iteration (SCI). The case of a single external coupling is illustrated by the Linac Coherent Light Source RF-gun cavity (Figure 3). The Omega3P computation requires two hours on 256 CPUs of an IBM SP3 computer at the National Energy Research Scientific Computing Center. The computation requires 400 gigabytes of memory, and the quadratic eigenvalue problem has about 3.2 million degrees of freedom. Lee presented simulation results for the international linear collider (Figure 4) with waveguide coupling of two types. To improve the solver efficiency, he pointed out, more research is needed to clarify the underlying structure of onlinear eigenvalue problems of this type.
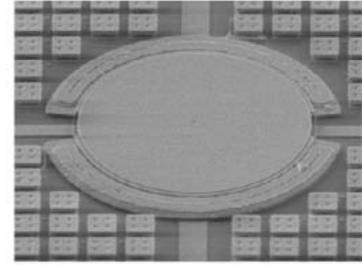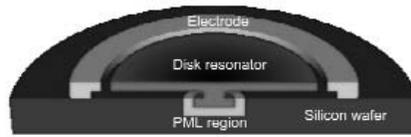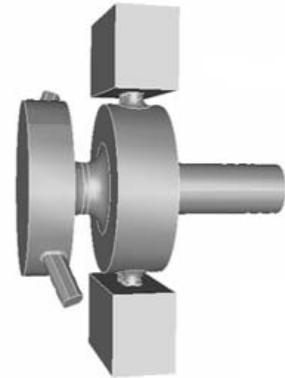

Figure 3. *Omega3P model of the Linac Coherent Light Source radiofrequency cavity. Courtesy of the Advanced Computation Department, SLAC.*

Figure 4. *Partitioned mesh of the International Linear Collider superstructure, which consists of two accelerator cavities with power and higher-order-mode couplers. Courtesy of the Advanced Computation Department, SLAC.*

The minisymposium on nonlinear eigenvalue problems elicited great interest and enthusiasm among both numerical analysts and application scientists at the SIAM meeting. It was evident from the talks that the rapid growth in computing power and the advances in numerical algorithms have enabled application scientists and industrial practitioners to introduce additional complexity in their computational models to account for more complex and nonlinear phenomena.

As computational tools for solving linear eigenvalue problems become mature, the research focus of the eigenvalue computation community is gradually shifting toward the development of methodologies for dealing with nonlinearity. Among the challenges confronting numerical analysts are not only standard convergence issues, but also the need for new algorithms that are capable of preserving structures of the eigenvalue problems.

Zhaojun Bai, a professor in the Departments of Computer Science and Mathematics at the University of California, Davis, and Chao Yang, a staff scientist in the Computational Research Division at Lawrence Berkeley National Laboratory, organized the two-part minisymposium on nonlinear eigenvalue problems. The presentations can be found at http://www.cs.ucdavis.edu/~bai/SIAMAN05.