

Anasazi software for the numerical solution of large-scale eigenvalue problems

C. G. Baker and U. L. Hetmaniuk and R. B. Lehoucq and H. K. Thornquist
Sandia National Laboratories

Anasazi is a package within the Trilinos software project that provides a framework for the iterative, numerical solution of large-scale eigenvalue problems. Anasazi is written in ANSI C++ and uses modern software paradigms to enable the research and development of eigensolver algorithms. Furthermore, Anasazi provides implementations for some of the most recent eigensolver methods. The purpose of our paper is to describe the design and development of the Anasazi framework. A performance comparison of Anasazi and the popular FORTRAN 77 code ARPACK are given.

Categories and Subject Descriptors: G.1.3 [Numerical Analysis]: Numerical Linear Algebra; G.4 [Mathematical Software]: ; D.2.13 [Software Engineering]: Reusable Software

General Terms: Algorithms, Design, Performance; Reliability, Theory

Additional Key Words and Phrases: Eigenvalue problems, Numerical Algorithms, Generic programming

Anasazi is a package within the Trilinos Project [Heroux et al. 2005] that uses ANSI C++ and modern software paradigms to implement algorithms for the numerical solution of large-scale eigenvalue problems. We define a large-scale eigenvalue problem to be one where a small number (relative to the dimension of the problem) of eigenvalues and the associated eigenspace are computed, and only knowledge of the underlying matrix via application on a vector (or group of vectors) is assumed.

An inspiration for Anasazi is the ARPACK [Lehoucq et al. 1998] FORTRAN 77 software library. ARPACK implements one algorithm, namely an implicitly restarted Arnoldi method [Sorensen 1992]. In contrast, Anasazi provides a software framework, including the necessary infrastructure, to implement a variety of algorithms. We justify our claims by implementing block variants of three popular algorithms: a Davidson [Morgan and Scott 1986] method, a Krylov-Schur [Stewart 2001a] method, and an implementation of LOBPCG [Knyazev 2001].

ARPACK has proven to be a popular and successful FORTRAN 77 library for

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000. Authors' addresses: C. G. Baker, U. L. Hetmaniuk, R. B. Lehoucq, Sandia National Laboratories, Computational Mathematics & Algorithms, MS 1320, P.O.Box 5800, Albuquerque, NM 87185-1320; email {cgbaker,ulhetma,rblehou}@sandia.gov. H. K. Thornquist, Sandia National Laboratories, Electrical and Microsystem Modeling, MS 0316, Albuquerque, NM 87185-0316; email hkthorn@sandia.gov.

Permission to make digital/hard copy of all or part of this material without fee for personal or classroom use provided that the copies are not made or distributed for profit or commercial advantage, the ACM copyright/server notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.

© 20YY ACM 0098-3500/20YY/1200-0001 \$5.00

the numerical solution of large-scale eigenvalue problems. A crucial reason for the popularity of ARPACK is the use of a reverse communication [Lehoucq et al. 1998, p. 3] interface for applying the necessary matrix-vector products. This allows ARPACK to provide a callback for the needed matrix-vector products in a simple fashion within FORTRAN 77. Unfortunately, the reverse communication interface is cumbersome, challenging to maintain, and does not allow data encapsulation. Moreover, because ARPACK uses a procedural programming paradigm where the matrix-vector operations rely upon the physical representation of the data manipulated, ARPACK is susceptible to design changes. Hence, code reuse is limited and software complexity and maintenance are more cumbersome.

The Anasazi framework employs more modern software development paradigms, both generic and object-oriented programming, via static and dynamic polymorphism [Vandevoorde and Josuttis 2002, Chapter 14], respectively. Static polymorphism, via templating of the linear algebra objects, allows algorithms in Anasazi to be written in a generic manner (i.e., independent of the data types). Dynamic polymorphism, via virtual functions and inheritance, allows eigensolvers to be decoupled from mechanisms such as orthogonalization and stopping conditions. Upshots of this decoupling are the facilitation of code reuse, increased algorithmic flexibility, and the ability to choose components at runtime.

We emphasize that our interest is not solely in modern software paradigms. Rather, our paper demonstrates that a rich collection of block eigensolvers is easily implemented using modern programming techniques. Our approach is algorithm-oriented [Musser and Stepanov 1994], in that requirements for efficient implementations of the necessary algorithms is considered first. This is then followed by a formulation of the software abstractions capable of implementing these algorithms, and their constituent mechanisms, in sufficiently diverse ways. The result is a collection of implementations that are efficient and flexible. We believe that Anasazi is the natural successor to ARPACK, inheriting and extending the quality practices employed by ARPACK.

There are related software efforts that implement several algorithms for solving large-scale eigenvalue problems (the reader is referred to [Hernández et al. 2005] for a software survey). The two most advanced, comparable software efforts are PRIMME and SLEPc:

- The Preconditioned Iterative Multi-Method Eigensolver (PRIMME) [Stathopoulos and McCombs 2006] is a C library for computing a number of eigenvalues and their corresponding eigenvectors of a real symmetric or complex Hermitian matrix. PRIMME provides a highly parametrized Jacobi-Davidson [Sleijpen and van der Vorst 1996] iteration, allowing the behavior of multiple eigensolvers to be obtained via the appropriate selection of parameters;
- The Scalable Library for Eigenvalue Problem Computations (SLEPc) [Hernández et al. 2006] library is another C library for the solution of large scale sparse eigenvalue problems on parallel computers. SLEPc is an extension of the popular PETSc [Balay et al. 2001] and can be used for either Hermitian or non-Hermitian, standard or generalized, eigenproblems.

PRIMME provides a flexible metasolver capable of implementing a variety of Hermitian eigensolvers. Predefined parameters are provided to emulate a large

number of powerful and popular eigensolvers, allowing easy use of the software by novice users. Expert users may manually specify the parameters in order to access the full flexibility available in the solver's behavior. Therefore, PRIMME is valuable both as a convenient eigensolver for practitioners and a platform for experimentation by eigensolver researchers. However, while parameters are provided to control mechanisms such as, e.g., stopping conditions and orthogonalization, the user is limited to the implementations provided by the developers of PRIMME. Furthermore, PRIMME provides implementations only over double precision real and complex fields. Support for `float` or extended precision scalar fields would require separate implementations due to the lack of generic programming ability in the C programming language.

SLEPc extends the PETSc toolkit to provide a library of solvers for standard or generalized, Hermitian or non-Hermitian eigenproblems. SLEPc provides wrappers for several eigensolver packages, most notably ARPACK and PRIMME, as well as native implementations of eigensolvers like Krylov-Schur, Arnoldi, and Lanczos. The use of PETSc also gives SLEPc users access to a large library of linear and nonlinear solvers, preconditioners and matrix formats. PETSc uses C language features such as `typedefs` and function pointers to support some generic programming and object-oriented paradigms. However, SLEPc's reliance on PETSc requires that the user employ PETSc for vectors and matrices. Similar to PRIMME, SLEPc can be compiled with support for double precision real or complex arithmetic, however only one version of the library can be used at a time. Furthermore, mechanisms such as orthogonalization are hard-coded allowing only parametrized control over their behavior.

The Anasazi framework was designed to include features from other eigensolver packages that are conducive to algorithm development, while avoiding some of the drawbacks mentioned above. The most important features that have been incorporated into its design are extensibility and interoperability. The extensibility of the Anasazi framework is demonstrated through the infrastructure's support for a significant class of large-scale eigenvalue algorithms. Extensions can be made through the addition of, or modification to, existing algorithms and auxiliary functionality such as orthogonalization, desired eigenvalue selection, and stopping conditions. This is encouraged by promoting code modularization and multiple levels of access to solvers and their data.

Interoperability in the Anasazi framework is enabled via the treatment of both matrices and vectors as opaque objects—only knowledge of the matrix and vectors via elementary operations is necessary. This permits algorithms to be implemented in a generic manner, requiring no knowledge of the underlying linear algebra types or their specific implementations. Furthermore, the Anasazi framework was designed to admit operation with any user choice of scalar field, vector and operator. This is accomplished using the template mechanism in the C++ programming language, an option not available to SLEPc or PRIMME. As a result, an Anasazi eigensolver using single-precision complex arithmetic can be used alongside another Anasazi eigensolver using an extended precision scalar type.

As a result of these design features, the Anasazi eigensolver framework is significantly more flexible than previous efforts, allowing its inclusion in diverse applica-

tion environments in addition to providing an arena for research into eigensolvers and their constituent mechanisms. The rest of this paper is organized as follows. Section 1 briefly discusses one class of algorithms that can be implemented using Anasazi in order to explore the types of operations necessary for an eigensolver framework. Section 2 reviews the Anasazi framework, discusses some of the design decisions, and illustrates the benefits of these decisions. Lastly, Section 3 provides some timings comparing ARPACK and Anasazi to demonstrate that object-oriented overhead has no impact on the performance of this modern software framework.

1. ALGORITHMIC INSPIRATION

The Anasazi software framework provides tools that are useful for solving a wide variety of eigenvalue problems. While development of the package continues to expand its scope to other important eigenvalue problems (e.g., nonlinear eigenproblems, constrained eigenproblems), the solvers currently released with the package are aimed at computing a partial eigen-decomposition for the generalized eigenvalue problem

$$\mathbf{A}\mathbf{x} = \mathbf{B}\mathbf{x}\lambda, \quad \mathbf{A}, \mathbf{B} \in \mathbb{C}^{n \times n}. \quad (1)$$

In this paper, the matrices \mathbf{A} and \mathbf{B} are large, possibly sparse, and we assume that only their application to a block of vectors is required. The reader is referred to [Saad 1992; Sorensen 2002; Stewart 2001b; van der Vorst 2002] for background information and references on the large-scale eigenvalue problem.

ALGORITHM 1.1: **Rayleigh-Ritz Algorithm**

- (1) *Let the matrix \mathbf{M} and bases \mathbf{U}, \mathbf{V} be given*
- (2) *Form the Rayleigh quotients $\mathbf{V}^H \mathbf{M} \Phi(\mathbf{A}) \mathbf{U}$ and $\mathbf{V}^H \mathbf{M} \Psi(\mathbf{B}) \mathbf{U}$ where $\Phi(\cdot)$ and $\Psi(\cdot)$ are matrix functions*
- (3) *Compute an eigen-decomposition (\mathbf{L}, \mathbf{W}) for the Rayleigh quotients $(\mathbf{V}^H \mathbf{M} \Phi(\mathbf{A}) \mathbf{U}) \mathbf{W} = (\mathbf{V}^H \mathbf{M} \Psi(\mathbf{B}) \mathbf{U}) \mathbf{W} \mathbf{L}$*
- (4) *Use the approximation (\mathbf{L}, \mathbf{W}) and the basis \mathbf{U} to construct an approximation for the pencil (\mathbf{A}, \mathbf{B})*

Algorithm 1.1 is a simple extension of the Rayleigh-Ritz procedure given in [Stewart 2001b, p.284]. This algorithm lists the salient steps found in the majority of large-scale eigensolvers, namely subspace projection methods. The matrices \mathbf{U} and \mathbf{V} are bases for the trial and test subspaces \mathcal{U} and \mathcal{V} , respectively. When these two subspaces are distinct, then the Rayleigh-Ritz method is called oblique. Otherwise, when $\mathcal{V} = \mathcal{U}$ the orthogonal Rayleigh-Ritz method results. The functions $\Phi(\cdot)$ and $\Psi(\cdot)$ are often used to improve convergence to the eigenvalues and eigenspace of interest. The matrix \mathbf{M} is often used to denote an inner product; for instance \mathbf{M} can be set equal to \mathbf{A} or \mathbf{B} when either matrix is Hermitian positive semi-definite.

The Rayleigh-Ritz procedure is useful for examining some of the functionality that should be provided by a general eigensolver framework. However, first we should observe some general structure of the participants in this algorithm. The bases \mathbf{U} and \mathbf{V} are dense matrices that are stored as a collection of vectors, which

we call a *multivector*. Interaction with \mathbf{A} , \mathbf{B} , and \mathbf{M} requires only knowledge of the underlying matrix via application on a vector or multivector, thus we regard these matrices as *operators*. The functions $\Phi(\cdot)$ and $\Psi(\cdot)$ should also be considered operators and, while the choice of functions is outside the scope of this paper, the use of these types of operators should not be hindered.

Given these observations, the functionality that is important to an eigensolver includes:

- multivector creation: create \mathbf{U} , \mathbf{V}
- operator-multivector applications: $\Phi(A)U$, $\Psi(B)U$;
- multivector arithmetic: $\mathbf{V}^H \mathbf{U}$;
- solution of typically much smaller eigenproblems (step 3).

A full list of our primitives for operators and multivectors will be presented in Section 2. It is worthy to note at this point that the Anasazi framework was designed to support block methods, defined as those that apply \mathbf{A} or \mathbf{B} to a collection of vectors, or multivector. One advantage of using a multivector data structure is that it can improve the ratio of floating-point operations to memory references and so better exploit a memory hierarchy.

ALGORITHM 1.2: Eigen-iteration

- (1) *Update the bases \mathbf{U} and \mathbf{V}*
- (2) *Determine whether any portion of the eigen-decomposition is of acceptable accuracy*
- (3) *Deflate the accurate portions of the eigen-decomposition*
- (4) *Terminate the eigen-iteration or return to step (1).*

Algorithm 1.1 needs to be augmented with several steps in order to result in an *eigen-iteration*. Algorithm 1.2 lists these additional steps and allow us to further analyze the components that make up an eigen-iteration. The decisions involved in Steps 2 and 4 require the determination of the interesting portion of the eigen-decomposition and a definition of accuracy or algorithmic breakdown, thus they are likely candidates for decoupling from the implementation of a particular eigen-iteration. For a particular eigen-iteration, deflation (Step 3) also may be conducted in a manner of ways. Steps 1 and 3, whatever their implementation, typically require orthogonalization methods, which are an active area of research. These methods provide another opportunity for decoupling functionality that need not be implemented in a specific manner.

This discussion illustrates that many distinct parts make up a large-scale eigensolver code: orthogonalization, sorting tools, dense linear algebra, convergence testing, multivector arithmetic, etc. Anasazi presents a framework of algorithmic components, decoupling operations where possible in order to simplify component verification, encourage code reuse, and maximize flexibility in implementation.

2. ANASAZI SOFTWARE FRAMEWORK

This section outlines the Anasazi software framework and discusses the design decisions made in the development of Anasazi. Three subsections describe the Anasazi operator/multivector interface, the eigensolver framework, and a review of the various classes in Anasazi. The reader is referred to [Baker et al. ; Sala et al. 2004] for software documentation and a tutorial.

We remark that Anasazi is largely independent of other Trilinos packages. Anasazi only relies on the Trilinos Teuchos package [Heroux et al.] that provides a common suite of tools, such as: `RCP`, a reference-counting smart pointer [Detlefs 1992; Bartlett 2004]; `ParameterList`, a list for algorithmic parameters of varying data types; and the BLAS [Lawson et al. 1979; Blackford et al. 2002] and LAPACK [Anderson et al. 1999] C++ wrappers.

2.1 The Anasazi Operator/Multivector Interface

Anasazi utilizes trait [Meyers 1995; Veldhuizen] classes to define interfaces for the scalar field, multivectors, and matrix operators. This allows generic programming techniques to be used when developing numerical algorithms in the Anasazi framework. Anasazi's eigensolver framework (Section 2.2) is comprised of abstract numerical interfaces that are all implemented using templates and the functionality of the template arguments is provided through their corresponding trait classes. Most classes in Anasazi accept three template parameters:

- a scalar type, describing the field over which the vectors and operators are defined;
- a multivector type, that depends upon the scalar type, providing a data structure that denotes a collection of vectors; and
- an operator type, that depends upon the multivector and scalar types, providing linear operators used to define eigenproblems and preconditioners.

Templating an eigensolver on operator, multivector, and scalar types makes software reuse easier. Consider in contrast that ARPACK implements the subroutines `SNAUPD`, `DNAUPD`, `CNAUPD`, and `ZNAUPD` for solving non-Hermitian eigenproblems. Separate subroutines are required for these four FORTRAN 77 floating point types (single and double precision real, and single and double precision complex). Moreover, four additional subroutines are needed for a distributed memory implementation. By templating abstract numerical interfaces on operator, multivector, and scalar types, it is only necessary to maintain a single code using the Anasazi framework.

Another aspect of software reuse that templating alleviates is through the separation of the eigensolver algorithm from the linear algebra data structures. This separation, as shown in Figure 2.1, allows a user of the Anasazi framework to leverage an existing linear algebra software investment. All that is required is the template instantiation of the trait classes, `MultiVecTraits` and `OperatorTraits`, for the user-defined multivector and operator, respectively. The `ScalarTraits` class and respective template instantiations for different scalar types are provided by the Trilinos Teuchos package [Heroux et al.]. Another user-friendly aspect of employing templates and traits mechanisms is that the Anasazi eigensolver, eigenproblem, and eigensolution are all defined by the specified scalar, multivector, and opera-

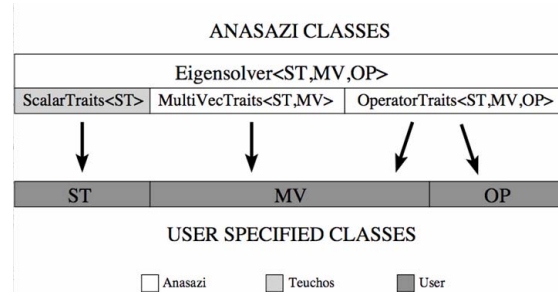


Fig. 1. An eigensolver templated on scalar (ST), multivector (MV), and operator (OP) type.

tor type at compile time. This approach, as opposed to using abstract interfaces and dynamic polymorphism, avoids any dynamic casting of the multivectors and operators in the user’s interaction with the Anasazi framework.

The `MultiVecTraits` and `OperatorTraits` classes specify the operations that the multivector and operator type must support in order for them to be used by Anasazi. Through the observations made in Section 1, it is clear that the `OperatorTraits` class only needs to provide one method, described in Table I, that applies an operator to a multivector. This interface defines the only interaction required from an operator, even though the underlying operator may be a matrix, spectral transformation, or preconditioner.

Table I. The method provided by the `OperatorTraits` interface.

OperatorTraits<ST,MV,OP>	
Method name	Description
Apply(A, X, Y)	Applies the operator A to the multivector X , placing the result in the multivector Y .

The methods defined by the `MultiVecTraits` class, listed in Table II, are the creational and arithmetic methods necessitated by the observations in Section 1. The creational methods generate empty or populated multivectors from a previously created multivector. The populated multivectors can be a deep copy, where the object contains the storage for the multivector entries, or a shallow copy, where the object has a view of another multivector’s storage. A shallow copy is useful when only a subset of the columns of a multivector is required for computation, which is a situation that commonly occurs during the generation of a Krylov subspace. All the creational methods return a reference-counted pointer [Detlefs 1992; Bartlett 2004] to the new multivector (`RCP<MV>`).

The arithmetic methods defined by the `MultiVecTraits` are essential to the computations required by the Rayleigh-Ritz method and the general eigen-iteration. The `MvTimesMatAddMv` and `MvAddMv` methods are necessary for updating the approximate eigenpairs in Step 4 of the Algorithm 1.1 or Step 2 of Algorithm 1.2. The `MvDot` and `MvTransMv` methods are required by the orthogonalization procedures utilized in Steps 1 and 3 of the eigen-iteration. The `MvScale` and `MvNorm` methods

Table II. The methods provided by the `MultiVecTraits` interface.

MultiVecTraits<ST,MV>	
<i>Method name</i>	<i>Description</i>
<code>Clone(X, numvecs)</code>	Creates a new multivector from X with $numvecs$ vectors.
<code>CloneCopy(X, index)</code>	Creates a new multivector with a copy of the contents of a subset of the multivector X (deep copy).
<code>CloneView(X, index)</code>	Creates a new multivector that shares the selected contents of a subset of the multivector X (shallow copy).
<code>GetVecLength(X)</code>	Returns the vector length of the multivector X .
<code>GetNumberVecs(X)</code>	Returns the number of vectors in the multivector X .
<code>MvTimesMatAddMv(alpha, X, M, beta, Y)</code>	Applies a dense matrix D to multivector X and accumulates the result into multivector Y : $Y \leftarrow \alpha XD + \beta Y$.
<code>MvAddMv(alpha, X, beta, Y)</code>	Performs multivector AXPBY: $Y \leftarrow \alpha X + \beta Y$.
<code>MvTransMv(alpha, X, Y, D)</code>	Computes the dense matrix $D \leftarrow \alpha X^H Y$.
<code>MvDot(X, Y, d)</code>	Computes the vector d where the components are the individual dot-products of the i -th columns of X and Y , i.e., $d[i] = X[i]^H Y[i]$.
<code>MvScale(X, d)</code>	Scales the i -th column of a multivector X by $d[i]$.
<code>MvNorm(X, d)</code>	Computes the 2-norm of each vector of X : $d[i] = \ X[i]\ _2$
<code>SetBlock(X, Y, index)</code>	Copies the vectors in X to a subset of vectors in Y .
<code>MvRandom(X)</code>	Replaces the entries in the multivector X with random numbers.
<code>MvInit(X, alpha)</code>	Replaces each entry in the multivector X with a scalar α .
<code>MvPrint(X)</code>	Prints the multivector X to an output stream.

are necessary, at the very least, for the computation of approximate eigenpairs and for some termination criteria (Step 4) of the eigen-iteration. Deflation and locking in Step 3 of the eigen-iteration necessitates the `SetBlock` method. Initialization of the bases for the eigen-iteration requires methods such as `MvRandom` and `MvInit`. The ability to perform error checking and debugging in Anasazi is supported by methods that give dimensional attributes (`GetVecLength`, `GetNumberVecs`) and allow the users to print out a multivector (`MvPrint`).

Specialization of the `MultiVecTraits` and `OperatorTraits` classes on given template arguments is compulsory for their usage in the eigensolver framework. However, Anasazi provides the following specializations of these trait classes:

- `Epetra_MultiVector` and `Epetra_Operator` (with scalar type `double`) allow Anasazi to be used with the Epetra [Heroux et al.] linear algebra library provided with Trilinos. This gives Anasazi the ability to interact with Trilinos packages that support the `Epetra_Operator` interface, like Amesos, AztecOO, Belos, Ifpack, ML, and NOX/LOCA.
- `Thyra::MultiVectorBase<ST>` and `Thyra::LinearOpBase<ST>` (with arbitrary scalar type `ST`) allow Anasazi to be used with any classes that implement the abstract interfaces provided by the Thyra [Bartlett et al.] package of Trilinos.

For scalar, multivector and operator types not covered by the provided specializations, alternative specializations of `MultiVecTraits` and `OperatorTraits` must be created. One benefit of the traits mechanism is that it does not require that

the data types are C++ classes. Furthermore, it does not require modification to existing data types, it only serves as a translator between the data type's native functionality and Anasazi's required functionality.

2.2 The Anasazi Eigensolver Framework

In this section we will discuss how to implement an eigensolver in Anasazi's framework. We will demonstrate Anasazi is a framework of algorithmic components, where decoupled operations simplify component verification, encourage code reuse, and maximize flexibility in implementation. This modularized approach requires a *solver manager* that combines a strategy with these algorithmic components to define an eigensolver. The high-level class collaboration graph for Anasazi's `SolverManager` class in Figure 2.2 lists all the algorithmic components offered by the Anasazi framework for implementing an eigensolver.

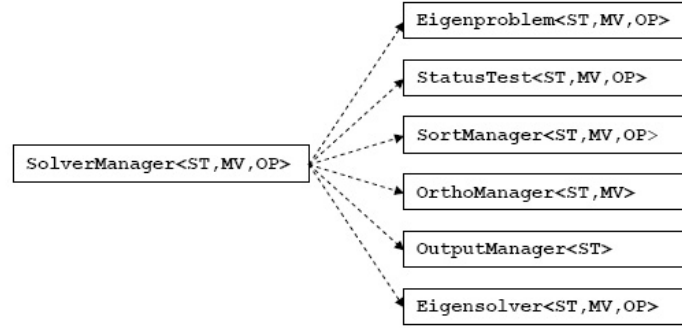


Fig. 2. Anasazi::SolverManager class collaboration graph.

The first component that is essential to the `SolverManager` class is the `Eigenproblem`. `Eigenproblem` is an abstract class that is a container for the components and solution of an eigenvalue problem. By requiring eigenproblems to derive from `Eigenproblem`, Anasazi defines a minimum interface that can be expected of all eigenvalue problems by the classes that will work with these problems. From a user's perspective, the most important part of the interface may be the methods for storing and retrieving the results of the eigenvalue computation:

```

const Eigensolution & Eigenproblem::getSolution();
void Eigenproblem::setSolution(const Eigensolution & sol);

```

The `Eigensolution` class was developed in order to facilitate setting and retrieving the solution data from an eigenproblem. Furthermore, the `Eigensolution` class was designed for storing solution data from both Hermitian and non-Hermitian eigenproblems. This structure contains the following information:

- RCP< MV > Evecs
The computed eigenvectors.

- RCP**< **MV** > **Espace**
An orthonormal basis for the computed eigenspace.
- std::vector**< **Value**< **ST** > > **Evals**
The computed eigenvalue approximations.
- std::vector**< **int** > **index**
An index into **Evecs** to enable compressed storage of eigenvectors for non-Hermitian problems.
- int numVecs**
The number of computed eigenpair approximations.

The **Value** structure is a simple container, templated on scalar type, that has two members: the real and imaginary part of an eigenvalue. The real and imaginary parts are stored as the magnitude type of the scalar type. The **Value** structure along with the **index** vector enable the **Eigensolution** structure to store the solutions from either real or complex, Hermitian or non-Hermitian eigenvalue problems. Implementations of the **SolverManager** class are expected to place the results of their computation in the **Eigenproblem** class using an **Eigensolution**.

The second component that is essential to a **SolverManager** class is the **Eigensolver**. The **Eigensolver** abstract base class defines the basic interface that must be met by any eigen-iteration class in Anasazi. Specific eigensolver iterations are implemented as derived classes of **Eigensolver**. This class defines two types of methods: status methods and solver-specific methods. A list of these methods is given in Table III. The status methods are defined by the **Eigensolver** abstract base class and rep-

Table III. A list of methods provided by any derived **Eigensolver**.

<i>Status Methods</i>	
<i>Method name</i>	<i>Description</i>
getNumIters	current number of iterations.
getRitzValues	most recent Ritz values.
getRitzVectors	most recent Ritz vectors.
getRitzIndex	Ritz index needed for indexing compressed Ritz vectors.
getResNorms	residual norms, with respect to the OrthoManager .
getRes2Norms	residual Euclidean norms.
getRitzRes2Norms	Ritz residual Euclidean norms.
getCurSubspaceDim	current subspace dimension.
getMaxSubspaceDim	maximum subspace dimension.
getBlockSize	block size.
<i>Solver-specific Methods</i>	
<i>Method name</i>	<i>Description</i>
getState	returns a specific structure with read-only pointers to the current state of the solver.
initialize	accepts a solver-specific structure enabling the user to initialize the solver with a particular state.

resent the information that any status test can request from any eigensolver. Each

eigensolver iteration also provides low-level, solver-specific methods for accessing and setting the state of the solver.

might need more information about solver state and compressed storage.

from the The combination of these two types of methods, along with the flexibility provided by status tests, provides the user with a large degree of control over eigensolver iterations.

explain why we desire status tests: runtime chosen stopping conditions, hook for checking state of solver for numerous other logic (restarting, debugging, etc). contrast against ARPACK.

The class `StatusTest` is used to specify stopping conditions for an eigen-iteration. `Eigensolver` queries the `StatusTest` during its class method `iterate()` to determine whether or not to continue iterating. Concrete subclasses of `StatusTest` provide particular stopping criteria. A typical interaction between these two classes is illustrated in Figure 3.

```
SomeEigensolver::iterate() {
    while ( somestatus.test.checkStatus(this) != Passed ) {
        //
        // perform eigensolver iterations
        //
    }
    return; // return back to caller
}
```

Fig. 3. Example of communication between status test and eigensolver

Each `StatusTest` provides a virtual method, `checkStatus()`, which queries the methods provided by `Eigensolver` and determines whether the solver meets the criteria defined by a particular status test. After a solver returns from `iterate()`, the caller has the ability to access the solver's state and the option to re-initialize the solver with a new state and continue iterating.

perhaps utility classes should be introduced before discussion of solver managers. emphasize: solver managers driven by parameter lists.

While this approach to interfacing with the solver is powerful, it can be overwhelming. It requires the user to construct a number of support classes and to manage calls to `Eigensolver::iterate()`. The `SolverManager` class was developed to encapsulate an instantiation of `Eigensolver`, providing additional functionality and handling low-level interaction with the eigensolver that a user may not want to specify. Solver managers are intended to be easy to use, while still providing the features and flexibility needed to solve large-scale eigenvalue problems.

For example, the constructor of `BlockDavidsonSolMgr` accepts only two arguments: an `Eigenproblem` specifying the eigenvalue problem to be solved and a `ParameterList` of options specific to this solver manager. This solver manager instantiates a `BlockDavidson` subclass of `Eigensolver`, along with the status tests and other support classes needed by the eigensolver, as specified by the parameter list. To solve the eigenvalue problem, the user simply calls the `solve()` method

of `BlockDavidsonSolMgr`. The solver manager calls `iterate()`, performs restarts and locking, and places the final solution into the `Eigenproblem`.

Under this framework, users have a number of options for performing eigenvalue computations with Anasazi:

- Use an existing solver manager. In this case, the user is limited to the functionality provided by the existing solver managers.
- Develop a new solver manager for an existing eigensolver. The user can extend the functionality provided by the eigensolver, specifying custom configurations for status tests, orthogonalization, restarting, locking, etc.
- Implement a new eigensolver (and so extend Anasazi). The user can write an eigensolver for an iteration that is not represented in Anasazi. The user still has the benefit of the support classes provided by Anasazi, and the knowledge that this effort can be easily employed by anyone already familiar with Anasazi.

2.3 Anasazi Classes

Anasazi is designed with extensibility in mind, so that users can augment the package with any special functionality that may be needed. However, the released version of Anasazi provides all functionality necessary for solving a wide variety of problems. This section lists and briefly describes the classes used in Anasazi.

Anasazi provides users with a concrete implementation of `Eigenproblem`, called `BasicEigenproblem`. This basic implementation provides all the functionality necessary to describe both generalized and standard, Hermitian and non-Hermitian linear eigenvalue problems.

However, a user working directly with an eigensolver (i.e., not with a solver manager) will need to recover the solution directly from the eigensolver state.

`SolverManager` defines only two methods: a constructor accepting an `Eigenproblem` and a parameter list of options specific to the solver manager; and a `solve()` method, taking no arguments and returning either `Converged` or `Unconverged` (Figure 4).

```
// create an eigenproblem
RCP< Anasazi::Eigenproblem<ST,MV,OP> > problem = ...;
// create a parameter list
ParameterList params;
params.set(...);
// create a solver manager
Anasazi::BlockDavidsonSolMgr<ST,MV,OP> solman(problem,params);
// solve the eigenvalue problem
Anasazi::ReturnType ret = solman.solve();
// get the solution from the problem
Anasazi::Eigensolution<ST,MV> sol = problem->getSolution();
```

Fig. 4. Sample code for solving an eigenvalue problem using a `SolverManager`

The goal of the solver manager is to instantiate a subclass of `Eigensolver`, along with the necessary support objects. Another purpose of many solver managers is to manage and initiate the repeated calls to the underlying solver's `iterate()`

method. For solvers that increase the dimension of trial and test subspaces (e.g., Davidson and Krylov subspace methods), the solver manager may also assume the task of restarting (so that storage costs may be fixed). This decoupling of restarting from the eigensolver is beneficial due to the numerous restarting techniques in use.

Performing an eigen-iteration requires a number of support classes. These are passed through the objects constructor, defined by **Eigensolver** to take the form listed in Figure 5.

```
Eigensolver(
    const RCP< Eigenproblem<ST,MV,OP> > &problem,
    const RCP< SortManager<ST,MV,OP> > &sorter,
    const RCP< OutputManager<ST> > &printer,
    const RCP< StatusTest<ST,MV,OP> > &tester,
    const RCP< OrthoManager<ST,OP> > &ortho,
    ParameterList &params
);
```

Fig. 5. Constructor for eigensolver

These support classes are employed for the following purposes:

- problem** - the eigenproblem to be solved; problem operators are defined.
- sorter** - the sort manager selects the eigenvalues of interest.
- printer** - the output manager dictates the verbosity level in addition to processing output streams.
- tester** - the status tester dictates the termination of the iteration `iterate()`.
- ortho** - the orthogonalization manager defines the inner product in addition to performing orthogonalization for the solver.
- params** - the parameter list specifies eigensolver-specific options.

The purpose of the **StatusTest** is to give the user or solver manager flexibility in terminating the eigensolver iterations in order to interact directly with the solver. For instance, typical reasons for terminating the iteration are:

- some convergence criterion has been satisfied;
- some portion of the subspace has reached sufficient accuracy to be deflated from the iterate or locked;
- the solver has performed a sufficient number of iterations.

The variation that exists for monitoring these and other conditions requires an abstract mechanism controlling the iteration.

The following is a list of Anasazi-provided status tests:

- StatusTestMaxIters** - monitors the number of iterations performed by the solver; it can be used to halt the solver at some maximum number of iterations or even to require some minimum number of iterations.
- StatusTestResNorm** - monitors the residual norms of the current iterate.
- StatusTestOrderedResNorm** - monitors the residual norms of the current iterate, but only considers the residuals associated with the most significant eigenvalues.

- StatusTestCombo** - a boolean combination of other status tests, creating near unlimited potential for complex status tests.
- StatusTestOutput** - a wrapper around another status test, allowing for printing of status information on a call to **checkStatus()**.

The purpose of a sort manager is to separate the eigensolver classes from the sorting functionality required by those classes. This satisfies the flexibility principle sought by Anasazi, by giving users the opportunity to perform the sorting in whatever manner is deemed to be most appropriate. Anasazi defines an abstract class **SortManager** with two methods, one for sorting real values and one for sorting complex values. Anasazi provides a concrete implementation called **BasicSort**. This class provides basic functionality for selecting significant eigenvalues: by largest or smallest real part, by largest or smallest imaginary part, or by largest or smallest magnitude.

Orthogonalization and orthonormalization are commonly performed computations in iterative eigensolvers. As explained in Section 1, all our current implementations are orthogonal Rayleigh-Ritz methods where an orthonormal basis representation is computed. The abstract base class **OrthoManager** defines a small number of orthogonalization-related operations, including choice of an inner product (e.g., Euclidean, induced by a symmetric positive semi-definite **B**). Combined with the plethora of available methods for performing these computations, Anasazi has left as much leeway to the users as possible. To this end, Anasazi provides two concrete orthogonalization managers:

- BasicOrthoManager** - performs orthogonalization using multiple steps of classical Gram-Schmidt [Daniel et al. 1976].
- SVQBOrthoManager** - performs orthogonalization using the SVQB orthogonalization technique described by Stathopoulos and Wu [Stathopoulos and Wu 2002].

In order to perform the Rayleigh-Ritz analysis used by the algorithms illustrating this section, Anasazi utilizes the classes **Teuchos::BLAS** and **Teuchos::LAPACK**. The purpose of these classes is to provide templated interfaces to the dense linear algebra routines provided by the BLAS and LAPACK libraries. Therefore, even such operations as dense matrix-matrix multiplication are made independent of the scalar field defining the eigenvalue problem. Users are therefore currently limited to algorithms provided by LAPACK.

3. BENCHMARKING

The benefits of an object-oriented eigensolver framework such as Anasazi are manifold: modularization provides improved code reuse, static polymorphism via templating allows easier code maintenance and a larger audience, and dynamic polymorphism via inheritance allows flexible runtime behavior. However, none of these benefits should come at the expense of code performance. Concern over overhead has long been an inhibiting factor in the adoption of object-oriented programming paradigms in scientific computing scenarios.

We now discuss the important issue of comparing Anasazi and ARPACK on a model problem. Our interest is in assessing any overhead of Anasazi and ARPACK, C++ and FORTRAN 77 software.

Table IV. Comparing the overhead of Anasazi with ARPACK; “—” denotes a measurement below the clock resolution.

Matrix size	Computing 50 Arnoldi vectors			
	Matrix-vector time [s]		Total runtime [s]	
	ARPACK	Anasazi	ARPACK	Anasazi
10000	—	0.01	0.14	0.15
62500	0.04	0.09	1.20	1.17
250000	0.15	0.32	4.98	4.79
1000000	0.66	1.23	19.2	18.8
Matrix size	Computing 100 Arnoldi vectors			
	Matrix-vector time [s]		Total runtime [s]	
	ARPACK	Anasazi	ARPACK	Anasazi
10000	0.03	0.02	0.53	0.55
62500	0.03	0.17	4.37	4.29
250000	0.34	0.64	17.8	17.5
1000000	1.27	2.40	68.4	67.1
Matrix size	Computing 150 Arnoldi vectors			
	Matrix-vector time [s]		Total runtime [s]	
	ARPACK	Anasazi	ARPACK	Anasazi
10000	0.03	0.04	1.15	1.22
62500	0.14	0.26	9.53	9.39
250000	0.50	0.96	38.1	38.0
1000000	1.97	3.56	149	146

We benchmarked Anasazi’s `BlockKrylovSchurSolMgr` (with a block size of one) and ARPACK’s `dnaupd` that compute approximations to the eigenspace of a non-symmetric matrix. Our goal was to benchmark the cost of computing 50, 100, 150 Arnoldi vectors for a finite difference approximation to a two dimensional convection diffusion problem. Both codes use the DGKS [Daniel et al. 1976] method for maintaining the numerical orthogonality of the Arnoldi basis vectors. The Intel 9.1 C++ and FORTRAN compilers were used with compiler switches “-O2 -xP” on an Intel Pentium D, 3GHz, 1MB L2 cache, 2GB main, Linux/FC5 PC.

rewrite this operator

The operator application in Anasazi records approximately twice as much time as the ARPACK implementation. This is because the Anasazi code used an Epetra sparse matrix representation, while the ARPACK implementation applies the block tridiagonal matrix via a stencil. Note that the operator application comprised only a small portion of the clock time in these tests. The performance of the Anasazi library in computing the Arnoldi vectors is similar to that of ARPACK. Our conclusion is that a well-designed library in C++ is as efficient as a FORTRAN 77 library.

4. CONCLUSION

reinforce their clarity

issues yet to be handled: anasazi provides only three eigensolvers, it also provides a framework capable of implementing multiple eigensolvers. for example, eigen-iterations requiring an iteration have been implemented using anasazi, such as RTR, TRACEMIN and Jacobi-Davidson.

5. ACKNOWLEDGMENTS

We thank Roscoe Bartlett, Mike Heroux, Roger Pawlowski, Eric Phipps, and Andy Salinger for many helpful discussions.

REFERENCES

- ANDERSON, E., BAI, Z., BISCHOF, C., BLACKFORD, S., DEMMEL, J., DONGARRA, J., CROZ, J. D., GREENBAUM, A., HAMMARLING, S., MCKENNEY, A., OSTROUCHOV, S., AND SORESENSEN, D. 1999. *LAPACK Users' Guide*, third ed. SIAM, Philadelphia.
- BAKER, C. G., HETMANIUK, U., LEHOUCQ, R. B., AND THORNQUIST, H. K. Anasazi: Block eigensolver package. See <http://software.sandia.gov/trilinos/packages/anasazi/index.html>.
- BALAY, S., BUSCHELMAN, K., GROPP, W. D., KAUSHIK, D., KNEPLEY, M. G., MCINNES, L. C., SMITH, B. F., AND ZHANG, H. 2001. PETSc Web page. <http://www.mcs.anl.gov/petsc>.
- BARTLETT, R., BOGGS, P., COFFEY, T., HEROUX, M., HOEKSTRA, R., HOWLE, V., LONG, K., PAWLOWSKI, R., PHIPPS, E., SPOTZ, B., THORNQUIST, H., AND WILLIAMS, A. Thyra: Interfaces for abstract numerical algorithms. See <http://software.sandia.gov/trilinos/packages/thyra/>.
- BARTLETT, R. A. 2004. Teuchos::RCP Beginner's Guide. Tech. Rep. SAND2004-3268, Sandia National Laboratories.
- BLACKFORD, L. S., DEMMEL, J., DONGARRA, J., DUFF, I., HAMMARLING, S., HENRY, G., HEROUX, M., KAUFMAN, L., LUMSDAINE, A., PETITET, A., POZO, R., REMINGTON, K., AND WHALEY, R. C. 2002. An updated set of Basic Linear Algebra Subprograms (BLAS). *ACM Transactions on Mathematical Software* 28, 2 (June), 135–151.
- DANIEL, J., GRAGG, W. B., KAUFMAN, L., AND STEWART, G. W. 1976. Reorthogonalization and stable algorithms for updating the Gram–Schmidt QR factorization. *Mathematics of Computation* 30, 772–795.
- DETLEFS, D. 1992. Garbage collection and run-time typing as a C++ library. In *Proceedings: USENIX C++ Technical Conference, August 10–13, 1992, Portland, OR*, USENIX, Ed. USENIX, pub-USENIX:adr, 37–56.
- HERNÁNDEZ, V., ROMÁN, J., TOMÁS, A., AND VIDAL, V. 2005. A survey of software for sparse eigenvalue problems. Tech. Rep. SLEPc Technical Report STR-6, Universidad Politecnica de Valencia. See <http://www.grycap.upv.es/slepc>.
- HERNÁNDEZ, V., ROMÁN, J., TOMÁS, A., AND VIDAL, V. 2006. SLEPc users manual: Scalable library for eigenvalue problem computations. Tech. Rep. DISC-II/24/02, Universidad Politecnica de Valencia. See <http://www.grycap.upv.es/slepc>.
- HEROUX, M., HOEKSTRA, R., SEXTON, P., SPOTZ, B., WILLENBRING, J., AND WILLIAMS, A. Epetra: Linear algebra services package. See <http://software.sandia.gov/trilinos/packages/epetra/>.
- HEROUX, M. A., BAKER, C. G., BARTLETT, R. A., KAMPSCHOFF, K., LONG, K. R., SEXTON, P. M., AND THORNQUIST, H. K. Teuchos: The trilinos tools library. See <http://software.sandia.gov/trilinos/packages/teuchos/>.
- HEROUX, M. A., BARTLETT, R. A., HOWLE, V. E., HOEKSTRA, R. J., HU, J. J., KOLDA, T. G., LEHOUCQ, R. B., LONG, K. R., PAWLOWSKI, R. P., PHIPPS, E. T., SALINGER, A. G., THORNQUIST, H. K., TUMINARO, R. S., WILLENBRING, J. M., WILLIAMS, A., AND STANLEY, K. S. 2005. An overview of the Trilinos project. *ACM Trans. Mathematical Software* 31, 3 (Sept.), 397–423.
- KNYAZEV, A. V. 2001. Toward the optimal preconditioned eigensolver: Locally optimal block preconditioned conjugate gradient method. *SIAM J. Scientific Computing* 23, 517–541.
- LAWSON, C. L., HANSON, R. J., KINCAID, D. R., AND KROGH, F. T. 1979. Basic Linear Algebra Subprograms for Fortran usage. *ACM Transactions on Mathematical Software* 5, 3 (Sept.), 308–323.
- LEHOUCQ, R. B., SORESENSEN, D. C., AND YANG, C. 1998. *ARPACK Users' Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*. SIAM, Philadelphia, PA.
- MEYERS, N. C. 1995. Traits: A new and useful template technique. *C++ Report* 7, 32–35.
- ACM Transactions on Mathematical Software, Vol. V, No. N, Month 20YY.

- MORGAN, R. B. AND SCOTT, D. S. 1986. Generalizations of davidson's method for computing eigenvalues of sparse symmetric matrices. *SIAM J. Scientific Computing* 7, 817–825.
- MUSSER, D. R. AND STEPANOV, A. A. 1994. Algorithm-oriented generic libraries. *Software Practice and Experience* 24, 7 (July), 623–642.
- SAAD, Y. 1992. *Numerical Methods for Large Eigenvalue Problems*. Halsted Press.
- SALA, M., HEROUX, M. A., AND DAY, D. M. 2004. Trilinos Tutorial. Tech. Rep. SAND2004-2189, Sandia National Laboratories.
- SLEIJPEN, G. L. G. AND VAN DER VORST, H. A. 1996. A Jacobi-Davidson iteration method for linear eigenvalue problems. *SIAM J. Matrix Analysis and Applications* 17, 2, 401–425.
- SORENSEN, D. 2002. *Numerical Methods for large eigenvalue problems*. Acta Numerica, vol. 11. Cambridge University Press, 519–584.
- SORENSEN, D. C. 1992. Implicit application of polynomial filters in a k -step Arnoldi method. *SIAM J. Matrix Analysis and Applications* 13, 357–385.
- STATHOPOULOS, A. AND MCCOMBS, J. R. 2006. PRIMME home page. See <http://www.cs.wm.edu/~andreas/software/>.
- STATHOPOULOS, A. AND WU, K. 2002. A block orthogonalization procedure with constant synchronization requirements. *SIAM J. Sci. Comput.* 23, 2165–2182.
- STEWART, G. W. 2001a. A Krylov-Schur algorithm for large eigenproblems. *SIAM J. Matrix Analysis and Applications* 23, 601–614.
- STEWART, G. W. 2001b. *Matrix Systems: Eigensystems*. Vol. II. SIAM.
- VAN DER VORST, H. A. 2002. Computational methods for large eigenvalue problems. P. Ciarlet and J. Lions, Eds. Handbook of Numerical Analysis, vol. VIII. North-Holland (Elsevier), Amsterdam, 3–179.
- VANDEVOORDE, D. AND JOSUTTIS, N. M. 2002. *C++ Templates*. Addison-Wesley Longman Publishing Co., Inc.
- VELDHUIZEN, T. Using C++ trait classes for scientific computing. See <http://oonumerics.org/blitz/traits.html>.