THE CURRENT STATUS OF UNSTEADY CFD APPROACHES FOR AERODYNAMIC FLOW CONTROL

Mark H. Carpenter * Bart A. Singer * Nail Yamaleev † Veer N. Vatsa * Sally A. Viken ‡ Harold L. Atkins *

1 Abstract

An overview of the current status of time dependent algorithms is presented. Special attention is given to algorithms used to predict fluid actuator flows, as well as other active and passive flow control devices. Capabilities for the next decade are predicted, and principal impediments to the progress of time-dependent algorithms are identified.

2 Introduction

Continuously expanding computer capabilities allow more attention to be devoted to the simulation of unsteady flows. At the turn of the millennium, practitioners routinely compute complex 3-D steady flows involving 10^5 - 10^6 grid points, and 2-D unsteady flows involving 10^5 - 10^6 points. These feats are performed while carrying five or more variables per node! If Moore's law persists (a fixed cost doubling of computer resources every 1.5 years) the next decade will provide practitioners with the resources to routinely simulate 3-D unsteady flows on 10^6 grid points. This computer capability will enable the burgeoning field of aerodynamic flow control (AFC), which is often time-dependent.

Active flow control offers the aerospace community the opportunity to expand the flight envelope through the use of steady suction/blowing, zero net mass synthetic jet actuators, or pulsed jets. These flow control devices exhibit promising flow control capabilities including separation control, thrust vectoring, mixing enhancement, noise control, and virtual shape change. Benefits of flow control include reduction in part-card count, empty weight, manufacturing costs, operating cost, fuel burn, and noise. A number of active flow control concepts have been tested in the laboratory and flight. Examples include leading edge suction for transition delay, zero net mass separation control and thrust vectoring fluidic injection. Computational studies have demonstrated that Reynolds averaged Navier Stokes (RANS) methodologies provide qualitative insight into active flow control applications. However, quantitative agreement is lacking between the computational and experimental results. To get from the bench top to real applications of flow control, reliable computational fluid dynamics (CFD) design tools must be developed and validated with the experimental and flight databases. An extensive amount of research is still needed to develop a production-type tool for active flow control applications for the design engineer.

A critical assessment of the current capabilities of time-dependent CFD, and identification of impediments that still exist is timely. We focus on identifying the critical areas (algorithmic and modeling) that possess notable leverage to the success of 3-D AFC computations.

The review of this material will be presented with the following strategy. Each section will begin with a broad overview of current state of the art in that field, followed by a description of general bottlenecks, and specific impediments for time dependent AFC computations. Finally, each section will conclude with a brief summary of NASA Langley Research Center's (LaRC) present research aimed at alleviating the bottlenecks, recognizing that some impediments are not being addressed due to limited resources. The fields of CFD and turbulence modeling are nearly boundless! To limit the scope of the review, only those methodologies which have shown promise in AFC simulations will be addressed.

The paper focuses on the general areas of algorithmic issues and turbulence models, and on the specific area of fluid actuators. The paper is organized as follows. Section 3 describes discretizations in space and time. The section begins with a broad overview of current state of the art in that field, followed by a description of general bottlenecks, and specific impediments for time dependent AFC computations. Finally, each section will conclude with a brief summary of NASA Langley Research Center's (LaRC) present research aimed at alleviating the bottlenecks, recognizing that some impediments are not being addressed due to limited resources. The fields of CFD and turbulence modeling are nearly boundless! To limit the scope of the review, only those methodologies which have shown promise in AFC simulations will be addressed.
tial discretizations. Section 4 describes algorithmic considerations related to convergence acceleration. Section 5 describes the current state of turbulence modeling for time-dependent flows. Section 6 describes specific considerations for effective actuator boundary conditions. Section 7 presents conclusions.

3 Discretizations: Time and Space

3.1 Why High-Order?

For reasons of efficiency, high-order schemes have long been advocated for use in time-dependent problems. In 1902, Kutta recognized the virtues of integrating ordinary differential equations (ODEs) with high-order schemes. The following simple example illustrates this point. Local error $\epsilon_i$ committed during one step of a temporal integration is described by the formula

$$ \| \epsilon_i \| \leq (\Delta t)^p \epsilon_i^{p+1} $$

where $p$ is the temporal order of the integration formula. Global error at time $T_f$ is estimated by summing all local errors after transporting each to the final time $T_f$. Estimates of global error, though not sharp, are generally expressed in the form

$$ \| E \| \leq (\Delta t)^p_{max} C(\exp[L(T_f - T_o)] - 1) $$

where $T_o$ is the initial time, and $C$ and $L$ are problem dependent constants related to solution smoothness, etc. In equation (2) the time-step satisfies $\Delta t < 1$, and a given error tolerance can be achieved by increasing the order $p$ while increasing the time-step. The resulting algorithm is more efficient if any additional work accrued at each large time-step, is more than compensated by a reduced number of steps. Work, however, increases with the order $p$. Near optimal values of $p$ in the range $3 < p < 5$ exist for countless stiff and non-stiff model problems. (see §IV.10 in Hairer and Wanner).

While it is generally recognized that high-order temporal schemes result in greater efficiency for time-dependent problems, it is less well appreciated that higher order spatial schemes additionally contribute to time-dependent efficiency! The virtues of high-order spatial schemes were first recognized and quantified by Kreiss and Oliger. To illustrate this advantage we provide an overview of the original argument presented in Kreiss and Oliger.

Consider the wave equation and initial data

$$ U_t + aU_x = 0, \quad U(x, 0) = e^{i k x} \quad (3) $$

on the space and time intervals $0 \leq x \leq 2\pi$ and $0 \leq t \leq T_f$, with the exact solution:

$$ U(x, t) = e^{i k (x - a t)} \quad (4) $$

Assume a uniform grid $x_j = j \Delta x$ with $\Delta x = 2\pi/N$. The general $n_t + n_x + 1$ point spatial discretization is

$$ U_x |_j = \frac{1}{\Delta x} \sum_{i=-n_t}^{n_x} a_i U(x_{j+i}), \quad j = 0, N - 1 \quad (5) $$

Substituting equation (5) into equation (3) and solving the system of ODEs ($U_t + MU = 0$) in Fourier space yields the modal solution

$$ \tilde{U}(x, t) = e^{i k (x - \tilde{a}(k) t)} \quad (6) $$

where $\tilde{a}(k)$ is the wave speed of the semi-discrete problem and is related to the Fourier image of the spatial operator. For example, the second and fourth order central difference waves speeds are

$$ a(2) = \frac{2 a \sin(k \Delta x)}{2 k \Delta x}, \quad a(4) = \frac{2 a \sin(k \Delta x) - \sin(2k \Delta x)}{12 k \Delta x} \quad (7) $$

For real $\tilde{a}(k)$, the difference (error) between the exact and numerical solutions (eqs. 4 and 6) obtained using trigonometric relations is $\epsilon(k) = 2 \sin(k[a - \tilde{a}(k)][t/2])$. Expanding the error in small phase angles yields the simple expression for the phase error:

$$ \epsilon(k) = | k [a - \tilde{a}(k)] t | \quad (9) $$

Taylor series arguments produce the leading order term for the difference in wave speeds

$$ a - \tilde{a}(k) \approx a \beta_p (k \Delta x)^p \quad (10) $$

with $\beta_p$ a scheme and order dependent constant. Substituting equation (10) into (9), defining the points per wavelength as $P = N/k$, and substituting $t = T_f$ yields

$$ \epsilon(k) \approx \beta_p k a T_f (\frac{2 \pi}{P})^p \quad (11) $$

The semi-discrete solution error accumulates linearly with time $T_f$, and is a strong function of the spatial truncation error. Rearranging equation (11) in terms of a maximum acceptable target error $\epsilon_T(k)$ yields the expression:

$$ P_p \geq 2 \pi \sqrt{\frac{\beta_p k a T_f}{\epsilon_T(k)}} \quad (12) $$

The grid points per wavelength necessary to achieve the specified target error $\epsilon_T(k)$ increases for problem size $2\pi$. The other dependencies rapidly decrease as the order of the spatial approximation is increased, motivating high-order spatial formulations.
in time-dependent problems. The cost of the computations increase with increasing order of accuracy, and a global minimum is reached for a finite value of $p$. Kreiss and Oliger\textsuperscript{63} suggest $4 \leq p \leq 6$ for problems of practical interest. Note that the optimal order for spatial and temporal operators is similar.

In summary, error accumulates linearly in time. The global error at $T_f$ is the sum of local errors that accumulate from each time-step in the integration. The local error at each time-step is the sum of three components: the temporal truncation error, the spatial truncation error, and the algebraic error. A simulation requiring many time-steps to reach $T_f$ requires extremely small local errors. High-order methods are the most efficient means of achieving these small local error tolerances.

An example will help to clarify this point. Consider a steady-state problem requiring lift to an engineering accuracy of three significant digits. A second order method could easily achieve this accuracy requirement. Now consider a similar time-dependent problem requiring lift (at the specified $T_f$) to an engineering accuracy of three significant digits. Further assume that 100 time-steps are required to integrate from $T_0$ to $T_f$. The local error (temporal, spatial, algebraic) at each time-step must be less than $10^{-5}$ to achieve the desired error tolerance of three significant digits. The constraint on spatial error ($10^{-5}$) in the time-dependent problem is much more severe than that required for the steady-state problem ($10^{-3}$). This example demonstrates the compelling need for high-order spatial operators especially for time-dependent simulations.

### 3.2 Temporal Algorithms

#### 3.2.1 Overview

The application of method of lines (MOL) to time-dependent partial differential equations (PDEs) results in an initial value problem (IVP) for a system of ODEs. Dozens of excellent texts with detailed descriptions of multi step, multi stage, and linear multi step methods have been written on the numerical integration of ODEs.\textsuperscript{26, 40, 55, 56, 86, 128} After more than 100 years of theoretical development, the mathematical framework for solving ODEs is relatively mature. In a general context, it is doubtful that dramatic (factors of 10) efficiency improvements can come from new methods.

The potential for dramatic efficiency improvements is greater in the field of time-dependent CFD, where current methodologies are surprisingly primitive. This schism between tidy mathematical theory, and rough CFD practices is not without good reason. Fluids practitioners are preoccupied with more urgent issues such as algebraic solvers, dimensionality issues, discontinuities, nonlinear instability, turbulence models, grid generation, etc. Nevertheless, the current objective is to identify mature technologies in the ODE literature that could have an immediate impact in CFD.

The hallmark of current ODE software is the ability to perform \textit{automated integration} for stiff ODEs. The first widely available multi step integration library was that developed by Gear,\textsuperscript{51} later modified and improved by Hindmarsh,\textsuperscript{50} resulting in the LSODE family of codes. Other variants have proliferated over the past two decades to account for the deficiencies of the original approaches (see VODE\textsuperscript{61}).

Automated integration begins with the user specifying 1) the system of ODEs, 2) the system Jacobian, and 3) the desired solution error tolerance. The software then automatically integrates the equations, using the most efficient numerical method chosen from a variety of candidate methods (second-order backward differentiation formulæ (BDF2) is frequently used). A reliable solution error estimator allows variable time-stepping. The time-step is adjusted to match the desired error tolerance. The resultant nonlinear system of algebraic equations is solved at each time-step using a Newton or modified Newton method. Direct matrix inversions are used within the Newton methods whenever possible. The algebraic error is reduced to a predetermined level, a constant multiple below the specified error tolerance. The Jacobian used in the nonlinear iteration is periodically reevaluated and stored based on the convergence rate of the iteration.

In contrast, the second-order accurate multi step BDF2 method is extensively used in the CFD community. System dimensionality prohibits the use of direct inverse methods useful for Newton or modified Newton methods. Iterative techniques such as Newton-Krylov methods are usually not as efficient as other more highly tuned methods (multigrid or combinations of methods). Error estimation or variable time-stepping mode is not perceived as necessary (correct or otherwise).

Three technologies presently used in the ODE community could have an immediate impact on CFD:

- high-order integrators: $p \geq 3$
- error estimation/variable time-stepping
- iteration termination strategies

To support this assertion, a brief summary of each area is presented.
3.2.2 High-Order Integrators

All general purpose solvers must integrate equations of considerable stiffness. We begin with a broad overview of stiffness, and identify the mathematical properties that enable a temporal integrator to efficiently integrate stiff equations.

Consider the integration of the system of ordinary differential equations represented by the equation

\[ \frac{dU}{dt} = S(U(t)) \]

In the present case, the vector \( S \) results from the semi-discretization (spatial and source terms) of the equations of fluid mechanics plus a suitable turbulence model. The integrator must integrate any \( S \) with which it is provided. Numerical difficulties often arise when the Jacobian of \( S, J = \partial S/\partial U \), has large eigenvalues. A useful definition for stiffness states that a problem is stiff when the largest eigenvalue of the Jacobian \( J \) (scaled by the time-step) \( ||z = \lambda(\Delta t)|| \) contained in the complex left-half plane \( (LHP) \) becomes much greater than unity. The resulting stiffness is then governed by both the Jacobian and the chosen time-step. Ideally, the time-step is selected solely based on error considerations and a good method simply executes this step-size in a stable and robust fashion. Time integration methods that do not amplify any LHP scaled eigenvalues are called \( \Lambda \)-stable. While \( \Lambda \)-stability is generally necessary, it is often not sufficient. We further demand that all eigenvalues \( ||z \rightarrow -\infty|| \) be completely damped. The combination of these two properties, \( \Lambda \)-stability and damping of \(-\infty\) eigenvalues, is defined as \( L \)-stability. General purpose solvers invariably rely on \( L \)-stable methods (and the partially stable \( L(\alpha) \) methods with suitable error controllers) to suppress temporal numerical instability and facilitate convergence of the nonlinear equation solver.

Popular implicit ODE integration methods are generally either distinctly multi step or multistage methods. Each has different strengths and weaknesses. Implicit multi step BDF methods compute each \( U \)-vector update to design order of accuracy using one nonlinear equation solve per step. Unfortunately, they are not \( \Lambda \)-stable above second-order. Additionally, they are not self-starting and have diminished stability properties when used in a variable step-size context. (Stability proofs are formulated assuming constant time-steps. Variable time-step cases may not be stable.) Practical experience indicates that large-scale engineering computations are seldom stable if run with BDF3.102 The BDF3 scheme, with its smaller regions of instability, is often stable but diverges for certain problems and some spatial operators. Thus, a conservative practitioner uses the BDF2 scheme exclusively for large scale computations due to its \( L \)-stability rather than \( L(\alpha) \)-stability.

Practical Runge-Kutta (RK) methods such as explicit, singly diagonal implicit, Runge-Kutta (ESDIRK) methods can be made arbitrarily high-order while retaining \( L \)-stability but possess intermediate \( U \)-vectors with a reduced order of accuracy and lesser stability. This reduced stage order may give rise to order reduction phenomena in the presence of substantial stiffness. ESDIRK schemes with \( s \) stages require \((s-1)\) nonlinear equation solves per step. Achieving progressively higher stage-order methods is possible with fully implicit methods such as the Radau IIA family. The cost of fully implicit methods greatly exceeds that of ESDIRK methods in the current context. Much less experience exists with implicit RK methods than BDF methods in the computation of large-scale engineering flows.

The general formula for a \( k \)-step, order-\( k \), BDF scheme can be written as

\[ U^{n+k} = -\sum_{i=0}^{k-1} a_i U^{n+i} + (\Delta t) b_k S^{n+k} \]

where \( n \) is the time-step index. At each time-step the BDF involve the storage of \( k + 1 \) levels of the solution vector \( U \), and the implicit solution of one set of nonlinear equations. Stability diagrams for these methods may be found in Hairer and Wanner.56 At order \( k > 2 \) an unstable zone for scaled eigenvalues in the complex LHP exists. At orders \( \{1, 2, 3, 4, 5, 6\} \) the methods are \( L(\alpha) \)-stable where \( \alpha \) is given by \( \{90^\circ, 90^\circ, 86.03^\circ, 73.35^\circ, 51.84^\circ, 17.84^\circ\} \). (The diffusion terms yield negative real eigenvalues. The convective terms discretized with nondissipative operators yield eigenvalues clustered on the imaginary axis \( \alpha = 90^\circ \). Numerical dissipation and boundary conditions displace both sets of eigenvalues into the LHP. Spatial operators with high levels of dissipation are more likely to be stable with BDF3.)

ESDIRK methods 78, 65 are implemented as

\[ U^k = U^n + (\Delta t) \sum_{j=1}^{k} a_{kj} S(U^j), \quad k = 1, s \]

\[ U^{n+1} = U^n + (\Delta t) \sum_{j=1}^{s} b_j S(U^j) \]

\[ \hat{U}^{n+1} = U^n + (\Delta t) \sum_{j=1}^{s} \hat{b}_j S(U^j) \]

where \( s \) is the number of stages, \( a_{kj} \) are the stage weights, \( b_j \) and \( \hat{b}_j \) are the main and embedded
scheme weights. The vectors U and \( \hat{U} \) are the \( p^{th} \) order and \( (p - 1)^{th} \) order solutions at time level \( n + 1 \). The vector \( \hat{U} \) is used solely for estimating error and is calculated at little extra cost. ESDIRK schemes differ from traditional SDIRK methods (see §IV.6 in Hairer and Wanner) by the choice \( a_{11} = 0 \), which permits stage-order two methods. The stiffly accurate assumption \( a_{ij} = b_j \) makes the new solution \( U_{n+1} \) independent of any implicit process within the integration step.

Many methods that combine multi step and multi stage schemes, generally referred to as linear multi step (LMS) methods, have been proposed in attempt to overcome the problems of each. Surprisingly, the results obtained with LMS methods (with a few exceptions) have been disappointing compared with either multi step or multi stage schemes. Notable LMS methods include the works of Cash.\(^{34,35}\) To increase the stability of the BDF methods, Cash proposed the extended backward differentiation formulae (EBDF) and the modified EBDF (MEBDF) schemes. The MEBDF schemes involve three stages to advance the solution one time-step. The first two stages are built from existing \( (p - 1)^{th} \) order BDF formulas, while the last stage combines the two previous BDF results into a \( p^{th} \) order solution. Note that the second BDF stage predicts a “super-future” point one time-step beyond the target time level, and substantially contributes to the \( \Lambda \)-stability of the method. At orders \( \{1, 2, 3, 4, 5, 6\} \) the methods are \( L(\alpha) \)-stable where \( \alpha \) is given by \{90°, 90°, 90°, 90°, 88.36°, 83.07°\}. The machinery involved with implementing the MEBDF algorithm is nearly identical to that involved in the BDF formulations. MEBDF schemes have the additional advantage that very accurate solution data are available on the first and third stages, based on previous information. This information can be used to provide the starting guess for the nonlinear iteration, and to establish time-step error estimates. The second stage typically uses the trivial guess as the starting point for the nonlinear iteration and no error estimate is made.

Butcher \(^{38}\) proposed a class of LMS methods for stiff differential equations. The new methods combine the properties of \( \Lambda \) and \( L \)-stability, and are reasonably simple to implement. They have a stability region that is identical to that of a RK method, but have high stage order. Uniformly high stage order eliminates the possibility of order reduction. The new methods were identified by focusing specifically on a diagonally implicit subclass of schemes referred to as DIMSIM.\(^{37}\)

3.2.3 Error Estimation

Temporal error management in the CFD community is presently accomplished by systematically halving the time-step until the solution is independent of further reduction. This strategy, while accomplishing the desired goal, can be streamlined by using an error estimator at each time-step and adjusting each time-step to attain the desired error.

Error estimation is accomplished by comparing two solutions of different orders \( U_{n+1} \) and \( U_{n+1}^{*} \) at the same time-step. For reasons of efficiency, the auxiliary solution \( U_{n+1}^{*} \) should be available at little additional cost. For example, in ESDIRK schemes (see eqn. 15), as well as MEBDF schemes, both \( U_{n+1} \) and \( U_{n+1}^{*} \) are constructed from available data. The difference \( ||U_{n+1} - U_{n+1}^{*}|| \) is proportional to the truncation error of the lower order formula \( U_{n+1}^{*} \). The estimate predicts the magnitude of the error in the solution, and gives insight into its overall quality. Frequently, linear and nonlinear instability can be predicted by the estimator well before the simulation diverges.

Figure (1) shows the error estimate (MEBDF4) for various \( \Delta t \). The test problem is for periodic shedding from the turbulent circular cylinder. The estimates are accurate to the correct order based on grid-converged data. The error estimate predicts that certain portions of the shedding cycle are more difficult to resolve in time. Variable time-stepping could easily increase the efficiency of the calculation by adjusting the time-step so that the same amount of error is produced at each time-step.

Variable time-stepping can introduce instability into some temporal integrators. The stability function of multistep schemes (BDF, MEBDF, LMS) is derived assuming constant time-steps. Large departures from constant step-size can lead to solution instability (although a good error estimator should warn this possibility). Conversely, the stability of multi stage schemes (ESDIRK) is independent of variable time-steps, because they are self-starting.

The time-steps in a variable time-step formulation, are chosen by a controller. A simple explicit controller is (see §IV.2 in Hairer and Wanner)

\[
(\Delta t)^{n+1} = (\Delta t)^{n} \left( \frac{Tol}{||U_{n+1} - \hat{U}_{n+1}||} \right)^{p}
\]

Similar yet more elaborate controllers exist for implicit formulations. The stability characteristics of a controller can be tuned/optimized in conjunction with the integration technique it is controlling. Together, they should meet the design objective and not introduce instability into the integration. Note that in figure (1) the predicted error
at finer tolerances has a high frequency component that the controller must suppress. (See Kennedy and Carpenter 7's for details on the feedback error controllers used with the ESDIRK scheme.)

3.2.4 Termination Strategy

An accurate error estimate can also be used to automate the termination strategy of the nonlinear iteration. Two competing components of temporal error are the truncation and algebraic errors. Truncation error is related to $\Delta t$ and the order of accuracy $p$, while algebraic error is the residual error generated each time-step by approximately solving the algebraic system. The local temporal error is the sum of the two components. To see full design order from the temporal scheme, the algebraic error must be driven below the truncation error at each time-step. This requires an accurate measure of truncation error, and must be provided by the error estimator.

The iteration termination strategy is complicated. Our experience indicates that design-order temporal convergence is achieved by maintaining a tolerance ratio of $10^{-2} < T < 10^{-1}$. Here $T$ is defined as the ratio of nonlinear algebraic error to temporal integration error at each time-step (or stage). Algebraic error for the nonlinear iteration is based on the $L_{\infty}$ norm of the density residual. Choosing the time-step based on accuracy considerations alone may not be the most efficient strategy for a temporal calculation. Decreasing the time-step can possibly greatly increase the convergence rate of the nonlinear algebraic system, thus increase efficiency. Gustafsson and Söderlind 54 devised optimal criteria for adjusting $\Delta t$. They assumed that either fixed point iterations, or modified Newton iterations is used for solving the algebraic system. The time-step is adjusted so that the iteration convergence rate approximately equals the optimal value. Because typical CFD algebraic solvers fall somewhere between fixed point and modified Newton iterations, additional work to refine these estimates is needed in the context of CFD time-dependent solvers.

3.2.5 Bottlenecks

Algebraic solvers that exhibit poor convergence behavior are an impediment for high-order schemes. Huge time-steps are needed to utilize the favorable aspects of high-order formulations. Algebraic solvers are needed that exhibits time-step independent convergence characteristics. If the convergence rate varies considerably with the time-step, then it may be more efficient to use a low-order scheme with small time-steps. Thus, high order temporal schemes need fast and robust algebraic solvers. Turbulent cases that have little or no convergence (one order of magnitude) present a second obstacle. A small number of cases are extremely difficult to converge, yielding dubious solutions at best. Nevertheless, solutions are still sought. It may be difficult to keep high-order formulations stable under these circumstances.

A perceptual impediment is the implementation of error estimation technology. A change in attitude about the nature of temporal error and the importance of its control is necessary. In spite of the perceived adequacy of existing temporal error practices, the CFD community should immediately adopt the practice of reporting a time-step error estimate as a necessary requirement of a high fidelity time-dependent simulation. Ideally, the estimate should include the component of primary interest in the simulation. For example, if lift and drag are the object of the study, then the estimate should include stepwise error estimates of these quantities, as well as information on which formulas were used to obtain the estimate.

Another common yet dangerous practice in the CFD community is to use a fixed number of iterations for each time-step. This approach eliminates the need for an iteration termination strategy and in most circumstances is satisfactory. Global temporal error [see equation (2)] strongly depends on time-steps with large local error. The error from just

![Figure 1. Time-dependent variation of predicted $L_2$ density error as calculated with the MEBDF4 scheme. The test case is turbulent flow around a circular cylinder, at $Re = 10^4$, and $Ma = 0.25$.](image-url)
one nonconvergent time-step potentially could dominate the error from all other time-steps combined! If an intrinsic feature of the flow significantly changes the convergence rate of the algebraic solver, then a fixed number of iterations is not a good strategy. The periodic blowing from zero mass fluid actuators is a prime example. Different phases of the cycle converge at different rates because convergence rate is sensitive to boundary conditions. A termination strategy that ensures a uniformly bounded algebraic error at each time-step is needed.

3.2.6 Langley effort

An ongoing effort focuses on the efficacy and efficiencies of several time integration schemes for the unsteady compressible Navier-Stokes equations. Existing and newly developed multistep and multi stage schemes are being studied, with particular attention to high-order \((p \geq 3)\) schemes. Past work includes comparisons of the high-order (ESDIRK4) Runge-Kutta scheme with first- and second-order BDF on laminar problems. Bijl, et al.\(^2\) showed that the efficiency of the ESDIRK4 scheme exceeds that of the BDF2 by a factor of \(2.5\) at engineering error tolerance levels \((10^{-1} - 10^{-2})\). Efficiency gains are more dramatic at smaller tolerances. No problems of nonlinear instability were noted with the high-order ESDIRK4 scheme on the problems tested.

Carpenter et al.\(^3\) has shown that stage order two Runge-Kutta schemes are susceptible to order reduction for stiff systems, although none is experienced for laminar problems with stiffness levels of \(O(10^3)\). However, turbulence models exhibit considerable stiffness at Reynolds numbers in the range of \(10^5 - 10^7\). Significant order reduction is experienced with ESDIRK4 for cases experiencing stiffness from strong turbulence fields. Ongoing studies include investigating the efficiency of ESDIRK schemes on other one- and two-equation turbulence models.

Figure 2 shows the convergence behavior of the ESDIRK4 scheme, the BDF2 and BDF3 schemes, and the MEBDF4 scheme. The test problem is the circular cylinder at Reynolds number \(10^4\), with a Mach number of 0.25. The calculations are run with the unstructured Fun2D code.\(^2\) Design order slopes are obtained for each scheme: 2, 3, 4, and 3 for BDF2, BDF3, MEBDF4, and ESDIRK4, respectively (note that we have accounted for the theoretical order of ESDIRK4 in accordance with order reduction). The MEBDF4 scheme was added to the comparison because it is a stage order three method and is not as susceptible to order reduction as the ESDIRK4 scheme. Research continues on establishing reliable error estimators and iteration termination strategies. A comparison of efficiency will be made between all integration schemes once each is automated.

A final observation is relevant to help focus future work. The BDF2 scheme yields engineering accuracy if each temporal mode in a time periodic flow is resolved with approximately 50 – 100 time-steps. The fourth-order ESDIRK formulation attains a similar accuracy (in spite of order reduction) using 5 – 10 time steps per period, with five stages per step, yielding approximately 25 – 50 time samples/period. State-of-the-art LMS methods could lower this estimate to 15 – 30 time-samples/period, an improvement of approximately \(O(10^{1/2})\) over existing temporal efficiency. The theoretical lower bound for temporal schemes, based on infinite order Chebyshev operators, is \(\pi\) samples/period. High-order schemes are presently asymptotically close to this theoretical lower bound. Another factor of three reduction in samples/period is perhaps all that remains and is becoming increasingly more difficult to attain. Algorithmic work focusing on other aspects of solver technology will have a greater chance of producing meaningful improvements during the next decade. The next section describes high-order spatial algorithms and their potential to increase the temporal efficiency, and section (4) describes the current and future status of algebraic solvers.
3.3 Spatial Algorithms

3.3.1 Overview

The spatial algorithms used currently in general purpose aerodynamics solvers have not changed appreciably during the past decade. Most current production codes (structured or unstructured) rely on some form of second-order upwind formulation with flux limiting to provide necessary robustness in the vicinity of unresolved features in the flow. Excellent texts describing these methodologies can be found elsewhere. For finite-difference methodologies see Hirsch,62,63, and LeVeque,88 For basic finite-element methodologies see Hughes,70, Zienkiewicz and Taylor,149,150, and Baker and Pepper.10

In section 3.1, we established that spatial algorithms play an important role in determining temporal efficiency. High-order methodologies will significantly contribute to the ultimate goal of efficient, general-purpose, time-dependent aerodynamic solvers. A broad overview of the spatial discretization landscape is now presented. Enabling technologies that allow extension of high-order methods into the general purpose aerodynamic solver arena are identified. A wealth of scientific literature supports the assertion that high-order general purpose algorithms will likely mature within the unstructured finite element framework within the next decade.

Spatial operators are categorized by the kind of grids on which they are formulated: structured or unstructured. A structured grid has large regions of the interior vertices that are topologically alike, which results in well-established connectivity patterns. Accommodation of complex geometries requires an arbitrary subdivision of the structured grid into what is referred to as a hybrid or multi block formulation. Three examples of structured codes currently used at Langley as general purpose aerodynamics solvers include the block structured TLNS3D,144 and CFL3D,142 and the overset-structured OVERFLOW.72 An unstructured mesh is one in which vertices may have arbitrarily varying local neighbors. Three examples of unstructured codes used at Langley (ICASE) are USM3D,49, FUN3D,2 and NSU3D.24 The distinction between structured and unstructured meshes usually (although not necessarily) extends to the shape of the elements: 2-D structured meshes typically use quadrilaterals, while unstructured meshes use triangles, with similar analogous element shapes in 3-D (hexahedra vs. tetrahedra).

Structured solvers offer simplicity, easy data access, and thus efficiency. The data structure and algorithmic simplicity of structured solvers leads to more efficiency and lower memory requirements for a given accuracy tolerance. A discrete derivative requires simple increments/decrements in array indices, in stark contrast to an unstructured formulation. The structured advantage in CPU time and memory can be as much as a factor of three on problems not requiring significant grid adaptation. However, on a complicated geometric domain a structured mesh may require many more elements than an unstructured mesh, because elements in a structured mesh cannot vary in size as rapidly. Structured grid generation approaches are far from being fully automated, and require user guidance in the decomposition step. A complicated 3-D structured mesh can take a month to generate. The current and future role of structured formulations is for repetitive computations late in the design cycle where grid templates might exist and grid generation and adaptation are not important components in the solution process.

Unstructured meshes offer flexibility in fitting complicated domains, rapid variation from small to large elements, and relative ease in refinement and de-refinement. Unlike structured mesh generation, unstructured mesh generation has been automated in mainstream computational geometry for some years. The major approaches for generating, refining, and improving unstructured meshes rely on unconstrained and constrained Delaunay triangulation, quad trees algorithms, or combinations of the above.19 A highly effective combination of techniques for high Reynolds number flows is an advancing layers method (ALM) 114 in the near-wall region, and an advancing front method (AFM) 91 in the far-field. Highly stretched viscous grids can be generated in a reasonably automated fashion with this approach. Automation begins to break down as aspect ratio increases on complex geometries.

Element shape has a profound impact on the accuracy and efficiency (direct and indirect) of a formulation. Meshes with unintended large aspect ratio cells lead to both poorly conditioned matrices and poor solution accuracy. Poor solution accuracy requires more grid points for a given accuracy. The additional cost of fixing a bad mesh can usually be mitigated by the faster convergence of the iterative solver.17 Babuška and Aziz17 showed that convergence on triangular elements is achieved only for angles bounded away from 180°. This rather weak condition becomes an issue for strongly anisotropic meshes used in high Reynolds number turbulent Navier-Stokes simulations. Near-wall aspect ratios on these grids can be in the range 10^4 – 10^6. Formulations typically try to limit the maximum angle in a grid (for example 179°) even though cur-
rent evidence is divided on the necessity of this condition. Quadrilateral and hexahedral meshes have an advantage in accuracy over triangular and tetrahedral meshes for these problems. The faces of hexahedral elements in the boundary layer are either almost parallel or almost orthogonal to the surface. Shock fronts and shear layers, which are also strongly anisotropic, require high aspect ratio cells for which the direction and location cannot be predicted in advance. Generation of these meshes can be difficult.

General purpose aerodynamic solvers have progressively shifted from hybrid/structured methods to unstructured formulations over the past decade. The principal motivations driving this change are grid generation on complex configurations and grid adaptation. These compelling reasons are likely to become more important during the next decade, particularly as the grid adaptation field matures for time-dependent simulations.

A large amount of inertia persists in the structured grid world, which is not entirely counterproductive. A time-dependent niche exists for computationally efficient formulations over the next decade. Unlike 3-D steady-state computations, realistic 3-D time-dependent computations are presently constrained by processor speed rather than memory requirements. The increased efficiency of structured methods is a notable advantage when run times can be decreased by a factor of two to three. The additional hybrid/structured grid generation time can be amortized if a calculation is likely to run for months.

3.3.2 High-Order Spatial Operators

High-order spatial operators need fewer points than second-order operators, to resolve the same information. The exact reduction strongly depends on the desired accuracy. Steady-state problems requiring an accuracy of three significant digits can be achieved with fourth-order schemes in half as many points in each spatial dimension. The total reduction in the number of points is approximately $O(10^4)$ in 3-D. Time dependent simulations that require solution accuracy to four or even five significant digits at each time-step, will favor high-order formulations to a larger degree. High-order spatial methods can increase the efficiency the time-dependent simulations by $O(10^3 - 10^2)$.

The constraints necessary to expedite grid generation and grid adaption will guide the next generation of high-order solvers. High-order methods must move beyond proof of concept and into the realm of being tools used to increase the efficiency of aerodynamic solvers.

The implementation of high-order methods is strongly dependent on whether the grid is structured or unstructured. High-order finite-elements (FE) are natural candidates for structured or unstructured meshes, while high-order finite-difference (FD) techniques are usually implemented on block-structured or overset grids. Finite-volume (FV) techniques exist in both forms; a close similarity between linear element FE methods and FV methods exists. All three approaches solve different forms of the governing integral equation. FD directly solves the integral equation by approximating the numerical fluxes. FV solves the divergence form of the integral equation by approximating the derivatives. FE take the divergence of the integral equations, multiply by an arbitrary test function, and integrate by parts. The solution itself is the resulting approximation.

Not all current general purpose spatial discretization algorithms are natural candidates for high-order extensions. For example, based on 2-D results, Casper and Atkins noted that a 3-D hybrid/structured essentially nonoscillatory (ENO)-FV formulation would be extremely expensive to implement relative to comparable FD techniques. Barth and Frederickson and Barth extended their unstructured FV solver to account for $k$-exact reconstruction. They note that comparing quadratic with linear reconstruction on a triangle requires roughly quadruple the number of solution unknowns. All high-order formulations require more work than second-order formulations, but some are more efficient than others.

Many different approaches to high-order FE have been adopted in developing numerical schemes to solve the compressible Euler equations. Two major classes have emerged as candidate schemes: 1) stabilized methods (continuous across interfaces), and 2) discontinuous methods (discontinuous across interfaces).

Standard Galerkin FE discretizations of convection-dominated Navier-Stokes equations produce wildly oscillating solutions unless dissipation terms are added to the formulation. Since the early 1980s stabilized FE methods have become increasingly popular in CFD. Early development motivated by the success of upwind FD/FV schemes included the streamline diffusion finite element Method (SDFEM), which later evolved into the streamline upwind Petrov-Galerkin (SUPG) scheme of Brooks and Hughes. A stabilizing term is added into the weak statement motivated by inviscid terms and results in a perturbed standard Galerkin test function. The stabilization creates an upwind effect by weighting more heavily the up-
stream nodes within each element. Hughes and Tezduyar 59 generalized the SUPG method to first-order hyperbolic systems and included work on the Euler equation. The original SUPG formulation suffered from oscillation in steep gradient regions (shocks) which led to the introduction of entropy variables and ultimately to stabilization terms including the effects of both the inviscid and viscous terms in their Galerkin least-squares (GLS) formulation.71 A variety of methodologies have been proposed to provide additional stability to the convection terms, monotone discrete solutions and ease of implementation.

Another approach that has been gaining in popularity in recent years is the discontinuous Galerkin (DG) method. The DG method originally introduced by Reed and Hill 117 exhibits several distinct advantages when applied to complex unstructured grids. Local polynomials are used to represent the data to arbitrary order, with the data on element interfaces treated as discontinuities. The approach is advantageous because the solution accuracy is relatively insensitive to mesh smoothness and can be extended to arbitrarily shaped elements. In 1988, Johnson and Pitkäranta 26 proved that the convergence rate of the method is $(\Delta x)^{k+1/2}$ for general triangulations. The method generates a local mass matrix that can easily be inverted, making the method efficient for explicit time integration. An entropy inequality for any scalar nonlinear equation 73 exists, implying discrete nonlinear $L_2$-stability for discontinuous solutions. (This assumes wellposedness and boundedness of the continuous nonlinear problem.) Several researchers have demonstrated superconvergence with DG.57 Lowrie et al. 22 obtained convergence rates of $2p + 1$, and Hu and Atkins 67 showed that the dispersion of the DG method is governed by $2p + 1$ for polynomials of order $p$. The DG formulation produces a matrix with many dense but small sub-matrices weakly coupled to their neighbors. Algorithmically, this matrix structure excels in a parallel environment and has high cache efficiency. Atkins and Shu 6 developed a quadrature-free approach that allows the precomputation and storage of much of the algorithm, thereby increasing the efficiency.

In the early 1980s, the $p$- and the $hp$-FEM methods were introduced by Babuška, and Szabó 8. They showed that for elliptic problems exponential convergence could be achieved with the $hp$-FEM method. The degree of the approximating polynomial can vary by elements so both grid refinement and order refinement are used simultaneously to attack solution error. The methods show design order for $p$ fixed in the limit $h \to 0$, and convergence for $h$ fixed $p \to \infty$. The behavior for high Reynolds number Navier-Stokes equations is less clear; nevertheless the $hp$-FEM methods have great potential in the context of complex geometries and grid adaptation.

High-order FD excel in their simplicity and efficiency, and in the richness of linear and nonlinear algorithmic permutations that can be formulated. The Achilles heels of FD are “boundaries” and “stability”.

Achieving numerical stability near boundaries with high-order FD stencils is difficult. This instability is closely related to the classical Runge oscillations exhibited by high-order polynomials on uniform grids near the boundaries. The solution in both cases is to lower the polynomial order, compress the grid, or increase the stencil width. Gustafsson 53 showed that to maintain spatial design order accuracy, the boundary stencil order must not deviate by more than one from the interior order of accuracy: fourth-order interior stencils require boundary closures of at least third-order accuracy! Note that in fluid dynamics applications, near-wall regions are precisely the regions where high-order accuracy is desirable. In general, low order treatments for boundary closures are not an acceptable alternative unless special near-wall grid refinement is used to compensate for reduced accuracy.

Strand 137 following the work of Kreiss and Scherer, 84 partially resolved the boundary closure dilemma by presenting constructive procedures for developing stable and accurate boundary schemes. Stability is ensured in an $L_2$ norm using a discrete summation-by-parts (SBP) procedure. Carpenter et al. 30 and Olsson 109, 110 showed how to impose the physical boundary conditions to preserve the SBP energy estimate. To date, boundary closures have been formulated for central and upwind FD schemes and Hermitian compact FD. FD schemes require smooth structured meshes that are often difficult to generate on complex geometries. Multi block/overset grids relax the gridding constraints, allowing piecewise smooth grids around complex geometries, but create a new set of complications. Conservation is a major concern on multi block grids and is extremely difficult to achieve on overset grids. Shocks and other discontinuities near interfaces must be treated carefully to ensure correct shock speeds and locations. In addition, interfaces, like boundaries, can cause linear and nonlinear instability and lead to decreased levels of solver robustness.

Multiple attempts have been made to overcome the difficulties of complex geometries for high-order FD schemes. By far, the most common solution to
ensure conservation and stability has been to reduce boundary or interface accuracy. As a general rule, this approach works well if little structure exists near the boundary or interface. Solution accuracy in complicated flow scenarios, however, is difficult to predict. Carpenter, Nordstrom, and Gottlieb, have developed $L_2$-stable interface conditions based on SBP energy estimates. The interface points are treated discontinuously through a penalty term and provide conservative, high-order solutions on multi block grids. The only grid requirement is $C_0$ interface continuity between blocks, a mild restriction.

### 3.3.3 Bottlenecks

The major obstacle facing all high-order spatial discretization methods FE or FD, structured or unstructured, is nonlinear instability in the presence of unresolved features. Shock and slippage are notable examples. The classical approaches to deal with discontinuities in FD and FD are the addition of local artificial viscosity and/or filtering, and total variation diminishing (TVD) or limiting approaches. Equivalent approaches exist in FE, though they are termed “stabilization.” The amount of added dissipation depends on the simulation objectives. Monotone solutions can be obtained with any formulation at the expense of reduced accuracy. In principle, the minimum amount of dissipation necessary for non-linear stability is advisable. Unfortunately, precise mathematical theory does not exist to determine the optimal dissipation. Thus, most approaches reduce the approximation/flux/solution near the discontinuity to first order to achieve monotonicity and robustness. Reduction to first-order accuracy locally results in the undesirable second-order global accuracy if a uniform $h$-refinement is then performed.

Artificial dissipation/filtering approaches are quite efficient and simple to implement in FD formulations. Unfortunately, they are often problem dependent, vary considerably on shock strengths, and are often user dependent. Although TVD and flux limiting approaches maintain monotonicity near discontinuities, they unfortunately degenerate to first-order accuracy near smooth extrema. (See LeVeque for an overview of TVD techniques.)

Essentially non oscillatory (ENO) and later Weighted ENO (WENO) schemes were developed to circumvent nonlinear instability. ENO schemes choose the smoothest stencil from all available design order stencils, thereby avoiding as much as possible interpolation/differentiation across discontinuities. ENO schemes have been extremely successful algorithms over the past 10 years for problems where both discontinuities and features requiring high-order spatial accuracy are required. See Shu for a detailed account of ENO/WENO schemes and their applications. The principal difficulty with structured grid ENO schemes is their extension to complex geometries. The mathematical foundations for ENO/WENO schemes (bounded total variation proofs, etc.) are predominantly based on periodic or infinite domains, and are outside the context of boundaries. ENO/WENO schemes can not be implemented at several points next to boundaries because they do not have smooth data outside the boundary to build high-order non-oscillatory stencils. The extension of ENO/WENO schemes to multiple domains is complicated by numerous boundary interfaces throughout the domain. Another difficulty with ENO schemes is that stencil searching algorithms have stencils that “switch” sometimes arbitrarily, which makes convergence to steady-state difficult. Atkins proposed a smoothly varying stencil biasing technique to eliminate this problem. Integer stencil shifts were not allowed in the approach. In addition, WENO schemes that are a smooth weighted sum of stencils in principle should not suffer from this difficulty.

Duruflsky, et al. and Abgrall attempted to overcome the geometric complexity problems by building fully unstructured stencils using stencil searching algorithms. Ollivier-Gooch suggested using a least-squares reconstruction approach for unstructured mesh ENO. Encouraging results were obtained for AGARD test case I with second- through fourth-order ENO, although convergence problems were experienced in the fourth-order case.

The use of overset meshes is another approach used to overcome geometric difficulties. Wang and Huang developed a compact ENO scheme and applied it to a multi domain overset code. Boundary issues with the ENO formulation are still present, and additional complications of non conservation at the interfaces exist. Wang et al. partially address the issue of interface conservation, but significant issues still remain for time-dependent discontinuous flows.

Most high-order formulations are only guaranteed stable on linear problems, and some cannot even claim linear stability. The high-order DG FE formulations can claim a stronger form of stability. Barth and Barth and Chirrier have designed numerical fluxes that satisfy a nonlinear energy condition. They assume a convex entropy extension of the Euler equations and bound the nonlinear “energy” of the system for all time in terms of the initial data. Simplified interface flux functions are derived to allow this result. These results and others provide an encouraging step towards nonlinearly stable for-
mulations that maintain high resolution.

Although theoretically advantageous, high-order spatial discretizations in their present form still have several obstacles to overcome. Considerable work has been devoted to these methods over the past decade, yet few studies demonstrate the increased efficiency of high-order spatial discretizations for general-geometry aerodynamic simulations, including turbulence models. De Rango and Zingg address this specific question and achieve encouraging results for which high order methods give more accurate solutions on a given grid. Additional work needs to be done to demonstrate increased efficiency for general geometry simulations, including turbulence models.

Geuzaine et al. and Delanaye et al. apply the quadratic reconstruction FV scheme of Barth and Frederickson to high Reynolds number flows. They achieve suitable convergence rates with generalized minimal residual (GMRES) and Bi-conjugate gradient stabilized (Bi-CGSTAB) algorithms, and show second- and third-order convergence on irregular and smooth meshes, respectively. They do not compare the efficiency of second- and third-order formulations, but note that the quadratic method converges more slowly than the comparable linear reconstruction. Delanaye and Liu report significant improvements in efficiency and accuracy in viscous flows, comparing the quadratic and linear formulations. Results in three dimensions are not as dramatic.

3.3.4 Langley Effort

Langley has had a strong presence over the last decade in the following high-order spatial disciplines: 1) structured grid FD, 2) structured grid ENO-FD and ENO-FV, 3) unstructured grid DG-FEM, and 4) unstructured grid SUPG-FEM. The following information is presented in the summary report to provide guidance when comparing the different high-order methods.

Table (1) can be used to compare the important attributes of current high-order formulations. Categories are rated on a scale from one to five, with five being the best currently available, and one being a capability representative of 1980. The categories are 1) complex geometry, 2) grid adaptation, 3) robustness (nonlinear), and 4) cost for a given accuracy requirement. At some level, all categories are closely related; however, we assume that each is independent from all others when assigning a numerical value. Specifically, the complex geometry category rates the capability of each method to accommodate complex 3-D configurations. This category is closely related to the locality of the discrete scheme. Grid adaptation is used to attack solution error. The second category, grid adaptation, describes each method's success on adapted grids, including sensitivity to grid smoothness and ease of grid generation. The robustness category describes the robustness of the method for under-resolved features and discontinuities. In simple terms, this category rate whether the code "runs" (converges for steady-state cases, and does not diverge in time-dependent cases) with minimal user support. The cost category describes the cost of achieving a given accuracy. It is assumed that the necessary grid has been generated by whatever means are necessary in each case.

The candidate schemes include two broad classes: 1) unstructured database schemes, and 2) semi-structured databases. We include the high-order FEM methods as unstructured methods because the structure within each element does not present a significant burden on the flexibility of the method. The unstructured candidate schemes are DG, SUPG, k-exact finite-volume, and k-exact ENO FV. The candidate structured grid schemes are Upwind FD, Upwind FV, and WENO-FD. The numbers are subjective, and should only be used as a relative guide for the purpose of comparing strengths and weaknesses. In general, researchers hold strongly varying opinions about the relative merits of each scheme.

Table 1: Comparison of high-order schemes. Categories are 1) complex geometry, 2) grid adaptation, 3) nonlinear robustness, and 4) cost for a given accuracy.

<table>
<thead>
<tr>
<th>Method</th>
<th>Geometry</th>
<th>Adapt</th>
<th>Robust</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstructured</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DG FE</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>SUPG FE</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>k-exact FV</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>LS-ENO FV</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Multi-block</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upwind FD</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Upwind FV</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>WENO-FD</td>
<td>2</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

4 Convergence Acceleration

A second algorithmic issue that contributes significantly to the efficiency of the temporal algorithm is the convergence rate of the algebraic solution algorithm. At some point in the solution process the algebraic system

$$A \tilde{x} = \tilde{b}$$  \hspace{1cm} (16)
is solved for the vector of data \( \bar{\varepsilon} \), with
\[
\mathbf{A} = \frac{\mathbf{I}}{\Delta t} + \partial \mathbf{F} \quad \quad (17)
\]
\[
\bar{\varepsilon} = \bar{U}^{k+1} - \bar{U}^k; \quad \hat{\bar{\varepsilon}} = \bar{R}(\bar{\varepsilon}) \quad \quad (18)
\]
The matrix \( \mathbf{A} \) is sparse. Equation (16) can be solved either directly, or iteratively. The dimensionality of \( \bar{\varepsilon} \) is \( O(10^7) \) for 3-D problems, which makes direct methods uncompetitive with iterative solvers. Shur-complement methods can be used to attack equation (16) by subdividing \( \mathbf{A} \) into numerous subproblems and interface conditions. Each subproblem is then solved directly, but the interface conditions are difficult (dense). (See Saad \(^{120} \) for details.)

Iterative methods fall into two broad categories: stationary and nonstationary.\(^{11} \) Stationary methods can be expressed in the simple form
\[
\bar{\varepsilon}^{k+1} = \mathbf{B}\bar{\varepsilon}^k + \bar{\zeta} \quad \quad (19)
\]
with the matrix \( \mathbf{B} \) independent of the iteration \( k \). These methods are older, simple to implement, and usually not very effective when used alone. Simple methods used in fluid mechanics are Jacobi, Gauss-Seidel (GS), symmetric GS, successive-over-relaxation (SOR), and symmetric SOR (SSOR). Numerous permutations of these methods exist including matrix reordering. (For details see Barrett et al.\(^{11} \)). More powerful stationary methods include incomplete factorization ILU(\(k\)), and block factorization. An ILU(\(k\)) approximately factors the original matrix \( \mathbf{A} \) using a LU decomposition with the level of fill governed by the parameter \( k \). Block factorization is motivated by tensor product grids (fine data dependencies) where implicit solves along directions are efficient. Both are somewhat more expensive than basic stationary methods, but have considerably faster rates of convergence. Implementation of sweeping algorithms is complicated in a parallel environment.

Nonstationary methods involve information that changes at every iteration. They are more recently developed, more difficult to understand, and more powerful. Commonly used examples in fluid mechanics include conjugate gradient (CG), generalized minimal residual (GMRES), bi-conjugate gradient (BiCG), quasi-minimal residual (QMR), and bi-conjugate gradient stabilized (Bi-CGSTab). These methods update the solution in certain "directions" by considering inner products of current residuals and other Krylov space vectors arising during the course of the iteration. (See Saad \(^{120} \) for details.) On different problem classes, the convergence rate of nonstationary methods varies considerably. Nachi-gal et al.\(^{104} \) showed that a class of problems exists for which each of the aforementioned nonstationary methods is a clear winner (in terms of efficiency), and a clear loser (in terms of efficiency).

A preconditioner is a matrix used to rotate a linear system into a new system that has the same solution but is easier to solve in some sense. A left-preconditioner acts on equation (16) yielding the new system,
\[
\mathbf{M}^{-1}\mathbf{A}\bar{\varepsilon} = \mathbf{M}^{-1}\hat{\bar{\varepsilon}} \quad \quad (20)
\]
where the new matrix \( \mathbf{M}^{-1}\mathbf{A} \) is easier to solve. The preconditioner changes the eigenstructure of the original system into a more compact set of eigenvalues that an iterative method can attack more effectively. Ideally, a preconditioner should change the eigenstructure dramatically but at a minimal additional cost. Simple preconditioners used in fluid dynamics include the block Jacobi, GS, and SSOR methods. More powerful preconditioners include incomplete factorization ILU(\(k\)) and block factorization.

Multigrid, at least in terms of elliptic problems, is a mechanism for rapid communication of multiscale information.\(^{29} \) Multigrid methods are usually defined as a strategy to accelerate any stationary or nonstationary iterative procedure. The solution is obtained on a sequence of grids, ranging from coarse to fine. Each grid smooths the high-frequency components of the residual on that grid. Restrictions and prolongations communicate the data between grids. The resulting algorithm rapidly communicates long wavelength data via the grids, while damping short wavelength data by using efficient local smoothing operators. The exact choice of grid structure, restriction and prolongation operators, and smoothers greatly influences the overall performance of the procedure.

Multigrid methods are effective techniques for accelerating convergence of elliptic and hyperbolic problems. Convergence rates easily approach 0.1 per cycle on elliptic problems such as the Poisson equation. The theoretical lower bound (if we rely on coarse grid corrections) on the convergence for hyperbolic equations is 0.75 per cycle for a second-order spatial discretization, and is routinely achieved by general-purpose Euler solvers. (See Molder\(^{103} \) for details.) The convergence rate suffers considerably for high-Reynolds number (turbulent) viscous flow solutions. The primary cause for this slowdown is the highly stretched wall normal grids used to resolve turbulent boundary layers. Wall normal stiffness can introduce stiffnesses of the order of \( 10^4 - 10^5 \). A secondary cause of slowdown is the existence of significant regions of low Mach num-
ber flow around stagnation points and in recirculation regions. A third cause of slowdown comes from the imperfections of realistic grids. Generating grids that have $10^7$ points without having grid anomalies is difficult. All general-purpose CFD solvers address each of these problems ultimately in the same way. For wall anisotropies, semi-coarsening and/or directional-implicit techniques precondition the wall normal boundary stiffness. Low Mach number preconditioning is added in those regions of the flow below a critical Mach number. Grid anomalies are addressed with grid smoothing and movement algorithms in problem regions.

Solution techniques within the aerodynamics community are far from being "black-box" algorithms. Two decades of experience have shown that none of these algorithms is well suited for solving broad classes of high Reynolds number turbulent flows. Practitioners rely on combinations of a wide variety of methods including 1) modified Newton-Krylov methods, 2) algebraic multigrid methods and 3) geometric multigrid methods, 4) defect-correction iteration techniques, and 5) sparse matrix methods.

Anderson et al.\textsuperscript{5} compared the efficiencies of several iterative strategies in the context of an unstructured, 3-D incompressible, Navier-Stokes solver. Multi element airfoils and high-lift systems were used as test problems in 2-D and 3-D. The turbulence model used was that of Spalart and Allmaras.\textsuperscript{132} GMRES was the Krylov method used in the study, with Gauss-Seidel or incomplete LU-decomposition used as a preconditioner. Newton-type solvers were shown to converge in the fewest iterations. In terms of work and storage the multigrid algorithms are the most effective means of reducing the residual on the problems studied.

In spite of all the powerful iterative techniques brought to bear on aerodynamic problems, the convergence rates for high-Reynolds problems can approach 0.98 per cycle. Mavriplis demonstrated the capabilities of his unstructured code NSU3D\textsuperscript{94} on complex 3-D configurations.\textsuperscript{95, 96, 97, 98, 99, 100} The features in NSU3D are among the most advanced presently used in the CFD community. A wide variety of 3-D test problems was run but not confined to 1) a realistic high-lift configuration including a wing, pylon, and nacelle, 2) the trapezoidal wing\textsuperscript{99}, and 3) an ONERA M6. Grids ranged from 1-10 million vertices. Reynolds numbers were in the 1-10 million range with wall-normal spacing of $10^{-3} - 10^{-5}$ based on chord length. Published convergence rates for these cases ranged from 0.96 - 0.98 per cycle. Approximately 500-1000 multigrid cycles were required to achieve residual levels converged to engineering tolerances.

4.1 Bottlenecks

The convergence rates quoted in the previous subsection are based on steady-state solvers. The efficiency of time-dependent and steady computations are closely related, as the underlying nonlinear matrices are nearly identical. (The time-dependent formulation converges slightly faster for time-steps governed solely by accuracy considerations). Thus, the cost of the unsteady calculations can be related to that of the steady. Simple back-of-the-envelope calculations reveal that at a minimum the unsteady computation will be equivalent to the solution of 100 steady-state problems each having the same computational complexity. Dropping an equation residual three orders of magnitude at a rate of 0.98 requires of approximately 400 iterations. Thus, modifying the convergence rate of the algebraic solution algorithm has a profound effect on the efficiency of the temporal algorithm. An order of magnitude increase in computational efficiency could be achieved by successful convergence acceleration efforts.

The convergence characteristics of high-order spatial formulations has not been extensively studied. The dissipation level of the spatial operator frequently affects the convergence rate of the algebraic system, with more dissipation producing faster convergence. High-order methods inherently have less dissipation, and could be more difficult to converge. An additional concern is the efficiency of multigrid methods as spatial order increases. The theoretical lower bound of multigrid methods (assuming coarse grid corrections) on linear advection is $\psi \geq 1 - 1/2^p$ where $\psi$ is the convergence rate, and $p$ is the order of spatial approximation. Fourth-order methods should converge no faster than 0.9375 per cycle. Ollivier-Gooch\textsuperscript{108} used multigrid on an unstructured high-order FV scheme and experienced increasing difficulty with convergence as the order of the approximation increased. Delaney et al.\textsuperscript{41} notes that quadratic elements converge more slowly than equivalent linear elements, when a difference GMRES algorithm is used. Interestingly, Bonhaus\textsuperscript{21} using SUPG with a Newton-Krylov GMRES method and diagonal preconditioning, experienced no changes in convergence rate with increasing order.

4.2 Langley Effort

Current iterative nonlinear solvers require $O(10^3)$ iterations to converge, based on high Reynolds number ($\approx 10^6 - 10^7$), turbulent, separated, 3-D, complex geometry flows. Newton solvers converge these problems in $O(10^1)$ iterations if a reasonable ini-
tial guess is given. Currently, a group at Langley is studying methods that potentially have textbook multigrid efficiency (TME). Methods that achieve TME converge at rates that are independent of the number of degrees of freedom (grid), and converge to the truncation error of the discretization in approximately $10^4$ work units. A work unit is defined as the work equivalent to the evaluation of the residual. Thus, TME methods have a potential increase in efficiency of $O(10^5)$ compared with existing state-of-the-art solvers. The potential efficiency of TME methods relies on the “factorizability” property of the Navier-Stokes (NS) equations. Splitting the NS equations into factors decouples the system somewhat, and allows optimal (fast and inexpensive) operators to be constructed for each portion. This “divide and conquer” approach yields nearly optimal efficiency for the entire problem. The Langley effort was showcased at the 2001 AIAA CFD conference. Work continues on this revolutionary method to implement TME algorithms on general purpose aerodynamic solvers.

5 Turbulence Modeling

5.1 Overview: Current practices

An equally important aspect of temporal algorithms is the underlying turbulence models being solved. Numerous turbulence models exist for attached steady turbulent flows. During the last two decades, great strides have been made in tuning these turbulence models to increase their robustness and generality. Practitioners routinely utilize these models to predict the flow behavior of a surprisingly broad class of complex problems with acceptable confidence levels in their solutions. Unfortunately, the same level of maturity does not exist for nonstationary flows, where time-averaged turbulence quantities inadequately describe the important dynamics of the flow. Oftentimes, these flows are dominated by large-scale flow features that are not properly modeled by conventional turbulence models. Flows with massive separations such as bluff-body wakes, cavity flows, shock-induced separations, and recirculation zones almost always fall into this category.

5.1.1 LES

In 1970 Deardorff published the first results of a large-eddy simulation (LES). The objective of an LES is to simulate, or directly compute, the large energy-containing fluid motions and to model only the small scales that are unresolved by the grid (the subgrid scales). A subgrid-scale (SGS) model acts to remove the energy associated with the small scales and therefore facilitates the global energy transfer from the large scales to the small scales. When properly implemented, LES can be used to simulate the turbulence in a flow at low to moderate Reynolds numbers. For more complete discussions of the concepts and applications of LES see Piomelli and the references therein.

In an LES calculation, the smallest resolved scales are determined by the grid-cell size. On finer grids, more of the flow is simulated and less is modeled. The SGS model therefore has an explicit dependence on the local grid-cell size. This feature of LES has led some to believe that the grid-cell size is unrestricted and simply corresponds to the break in resolved and subgrid scales. This sort of thinking usually leads to poor calculations. An important assumption in an LES calculation is that the energy-containing scales are actually simulated. To do this, the peak in the energy spectrum must be in the resolved range of scales and the cutoff between the resolved and subgrid scales should be in the inertial wavenumber range.

The development of dynamic subgrid-scale models for LES has motivated the use of LES on nonstationary turbulent flows over extremely complex configurations. However, estimates of the grid requirements for LES computations over realistic geometries at realistic Reynolds numbers indicate that LES is not likely to be a viable option for most flows for several decades to come.

5.1.2 URANS

Despite the abovementioned shortcomings of existing turbulence models, most time-dependent codes use unsteady extensions of popular steady-state algorithms. These calculations are typically referred to as unsteady Reynolds-averaged Navier Stokes (URANS). The turbulence models used in URANS span the spectrum of available models.

Höld et al. used a Baldwin-Lomax model to compute the unsteady flow about a rounded half-cylinder embedded in a flat plate. Their time-averaged pressure coefficients agreed reasonably well with experimental measurements. Surface pressure spectra, however, showed far more disagreement with the measured spectra. Höld et al. suggested that the disagreements were primarily temporal and spatial resolution issues rather than a turbulence modeling problem. Their paper illustrates a significant problem associated with assessing turbulence models for unsteady flows; unless care is taken to first assess grid and time-step issues, the influence of the turbulence model is almost impossible to evaluate. Because of the computational expense of doing thorough resolution checks for unsteady problems, these numerical issues are often left unresolved.
Khorrami, Berkman, and Choudhari\textsuperscript{79} and later Khorrami, Singer, and Berkman\textsuperscript{80} used the two-equation SST (k-\omega) model of Menter\textsuperscript{101} for 2-D unsteady calculations of flow in the vicinity of a leading-edge slat. In their work the grid resolution and time-step were sufficiently fine for them to suspect that the turbulence model was responsible for excessive damping of the large eddies in the slat cove. Subsequent work by Khorrami, Singer, and Lockard\textsuperscript{81} locally eliminated the turbulence production term in much of the slat cove and observed unsteady coherent vortex motions that closely resembled the PIV measurements reported in the work of others.\textsuperscript{112, 140, 141}

Scotti and Piomelli\textsuperscript{121} assessed four low Reynolds number turbulence models in pulsating channel flows generated by oscillating the imposed pressure gradient. The apparent simplicity of these flows was deceptive. In fact the pulsating flows represented significant challenges to the URANS models. The turbulence was out of equilibrium, in some cases so much so that partial relaminarization occurred, followed by a re-transition. Somewhat surprisingly, the phase-averaged streamwise velocity profiles matched corresponding LES results reasonably well when plotted in wall coordinates. However the time-averaged velocities of the URANS models were always less than those determined from the LES. Significant differences in the Reynolds shear stresses and the turbulent kinetic energies were also observed.

5.1.3 Composite LES/RANS Schemes

Recently, a class of composite models has been developed for unsteady flows. These composite models attempt to blend the unsteady capabilities of LES with a method having grid requirements that are more like those conventionally used in Reynolds-averaged Navier-Stokes (RANS) calculations. A composite model will typically involve a RANS turbulence model and RANS-type grid in regions near solid surfaces, where the resolution of turbulent eddies would require exceptionally fine grid resolution. In these regions typical RANS models are usually adequate for modeling the effects of turbulent flow. Away from wall regions, where large unsteady flow structures dominate the flow, the composite model shifts from a RANS-type turbulence model to a less-dissipative LES-type subgrid-scale model. Similarly, away from wall regions, the gridding strategy will shift from a RANS-type grid to a grid amenable to resolving at least some of the unsteady turbulence scales of motions.

A variety of different composite models have been proposed in the last few years. The composite models go by different names and involve different turbulence models for the RANS and LES regions and different switching functions to shift between the RANS and LES regions. Batten, Goldberg, and Chakravarthy\textsuperscript{18} proposed limited Navier Stokes; Zhang, Bachman, and Fasel\textsuperscript{138} developed their flow simulation methodology; Arunajatesan, Shipman, and Sinha\textsuperscript{14} developed the hybrid RANS-LES. However, the first and perhaps the most thoroughly tested of these composite models has been the detached eddy simulation (DES) model of Spalart.\textsuperscript{135} In its most common form, the DES model is implemented in combination with the Spalart-Allmaras (SA) turbulence model,\textsuperscript{132} although recently DES results with Menter’s SST model\textsuperscript{101} have also been published.\textsuperscript{138} In either case, the switch from the RANS region to the LES region is governed by a ratio of grid-cell size with distance from the wall. Viscous grid cells (high aspect ratio with lengths long compared to the distance from the wall) characterize the RANS regions. Small, approximately uniformly sized grid cells far from solid surfaces characterize the LES region. In theory, the RANS portion correctly models the rather benign attached turbulent flow regions, while the LES portion has sufficient grid resolution and sufficiently reduced eddy viscosities to simulate the large-scale turbulent flow structures.

The ease with which DES can be implemented in a code is a mixed blessing. Upgrading an existing SA model to accommodate the DES model requires minimal additional logic. The triviality of the modifications required for this transformation has led numerous well-intentioned users to perform and publish DES-like calculations. Unfortunately, it is easier to obtain results than to verify and critically interpret them.

Proper grid resolution is always important in CFD. However, with DES, the grid takes on new significance, as the grid determines the switch between the RANS region and the LES region. Fortunately, Spalart\textsuperscript{135} provides some guidance for the development of appropriate grids. Importantly, once the RANS region of the flow is adequately resolved, improved grid resolution should be undertaken local to the LES portion of the simulation, not globally. Over-resolving the RANS regions can displace the RANS/LES switch over surface into the boundary layer, where the grid is likely to be insufficient for a reliable LES calculation. Of course, continued grid refinement will eventually lead to adequate LES grids over the entire flow, although for aerodynamic flows of commercial interest, solving the flow on such a grid will likely be a decades-long endeavor.
With a DES grid well designed for a particular application, global grid refinement is likely to adversely affects the results. Nikitin et al.\textsuperscript{109} discusses these circumstances in the context of a series of channel-flow calculations. Although modifying the switching function, as is done in other composite LES/RANS methods, can alleviate some of the deleterious effects, a key to successful implementation of these methods is going to involve careful attention to gridding.

Another concern with the implementation of composite methods is the numerical diffusion, especially in the LES region where diffusion associated with the model is far less than in the RANS region. In simulations of the unsteady flow about a sphere, Constantinou and Squires\textsuperscript{38} used both second-order and fifth-order accurate upwind differences for the convective terms. All other operators were discretized using three-point, second-order accurate central differences. On the same grid, they found that the results of calculations with the fifth-order accurate scheme agreed better with a separate LES calculation than the results obtained with the second-order upwind scheme.

In later work, Strelets\textsuperscript{38} was concerned about the diffusion of upwind schemes in the LES region and the lack of stability of central-difference schemes in the RANS regions. To weaken the adverse effects of upwinding in the LES regions, he used a hybrid central/upwind approximation of the inviscid fluxes. The hybrid is designed so that in the RANS region the scheme is "almost upwind" and in the LES regions the scheme is "almost centered."

Recent work at NASA Langley suggests that second-order central-difference schemes can be used for both the RANS and LES regions, although the grid resolution must be chosen appropriately. The key is to ensure that the turbulence model, and not the numerics, controls the diffusion.

DES has been used in unstructured grid environments, first by Forsythe, Hoffmann, and Dietiker.\textsuperscript{48} Their research suggested that \( C_{DES} \), the one additional free parameter in DES (as compared to SA) needed to be adjusted when DES is used in conjunction with tetrahedrons. Because \( C_{DES} \) weights a measure of grid-cell size in the RANS/LES switching function, the need to adjust \( C_{DES} \) when differently shaped grid elements are used is not surprising. Pelaez, Mavriplis, and Kandil\textsuperscript{113} used the standard value of \( C_{DES} \) for their unstructured grid calculations and ultimately concluded that the value should be determined by performing a decaying homogeneous turbulence test case.

In spite of the many unanswered questions associated with DES, in the right hands it can be a useful tool for simulating unsteady complex flows. Travin et al.\textsuperscript{143} used DES for flows around a circular cylinder. They explored the effects of grid, Reynolds number, 2-D versus 3-D, DES versus URANS, and laminar versus turbulent separation. An important conclusion, first suggested by Spalart et al.\textsuperscript{134} is that 2-D DES calculations are not fruitful. To achieve the advantages of DES, a true LES should be done in the LES regions which implies that the full 3-D flow needs to be simulated. Without the third dimension, important turbulence dynamics are inoperative. Another important conclusion is that the benefits of DES vary with the flow and the information desired. In the case of a circular cylinder with laminar separation, DES proved itself superior to URANS, even for obtaining time-averaged quantities. However, for the case of turbulent separation, the time-averaged quantities were the same, regardless of whether DES or URANS was used. Hedges, Travin, and Spalart\textsuperscript{50} found that URANS performed almost as well as DES for time-averaged quantities on a four-wheel landing-gear model. The unsteady flow around the landing gear appeared much more turbulent-like in the case of DES compared to URANS, but unfortunately unsteady flow data was not available for comparison.

### 5.2 Bottlenecks

In this section we will review some of the bottlenecks that are holding back further progress in resolving turbulence modeling issues. In particular, we will discuss experimental validation problems, consistency concerns for the composite methods, and the high cost of the calculations.

#### 5.2.1 Validation by Experiments

One of the great difficulties with new turbulence models for unsteady flows is the lack of relevant experimental data with which to validate the models. As discussed earlier, different types of turbulence models with very different unsteady characteristics can provide very similar time-averaged results. Hence, if the time-dependent behavior of the flow is important in an application, then experiments that provide time-dependent data are indispensable.

Although flow past circular cylinders and spheres has provided useful validation tests in the past, for applications that involve flow actuators, particularly those that involve unsteady mass addition, new types of experiments are required.

A typical unsteady flow actuator comprises a nozzle or orifice that communicates with a cavity. The cavity includes either a flexible membrane or a moving wall. Unsteady deformation of the flexible mem-
ber inside the cavity results in a local pressure change that drives fluid through the nozzle and into an external flow. In practical applications, the external flow is turbulent without the presence of the actuator; the flow inside the actuator is often a mix of laminar and turbulent regions.

A good place to start with the experiments is where a number of calculations of flow actuators have begun, i.e., with actuators in no-flow and laminar flow environments. Such experiments are currently being performed at Langley. These experiments are time-consuming and require attention to details not typically recognized as important. For instance, the deformation of the flexible member may need to be simulated or modeled with greater fidelity than expected, time lags between the electrical actuation of the flexible member and its actual mechanical actuation need to be addressed, and flow measurements in largely inaccessible locations, like the actuator cavity, are desired.

The problems prove much more difficult when the external flow is turbulent. In that case, characteristics of the external turbulent boundary layer must be measured. In addition, the interaction of the flow structures in the turbulent boundary layer with the flow in the cavity must be characterized. Validated criteria for choosing a particular turbulence modeling approach for a specific type of actuator flow have not been established. However, we speculate that if a significant amount of the external flow is ingested into the actuator cavity, then a full LES calculation may be the only reasonable approach. On the other hand, if the actuator is biased such that it does not ingest the external flow, a DES calculation might be able to calculate the relevant flow features. One concern with DES in this sort of application is that the relevant unsteady flow regions are close to the wall; hence the model may revert to its RANS characteristics in these regions and be too diffusive. Other composite LES/RANS models may be more appropriate for such flows. Although explorations of these problems are ongoing at Langley, no definitive guidance can be provided to the CFD community yet.

5.2.2 Consistency Concerns

Gatski\textsuperscript{50} raises questions as to the formal consistency of the composite methods. The composite methods assume that the flow variables from the RANS regions match smoothly with the corresponding variables from the LES regions. Gatski\textsuperscript{50} shows that such a match cannot be taken for granted, even without considering the consistency issues associated with the coupling of the temporal averages used to derive RANS models and the spatial filtering used to derive SGS models for the LES regions. To date, these consistency issues have not been addressed by the modeling community.

5.2.3 Calculation Cost

One of the problems with using a composite LES/RANS method is that for the LES region to be reasonably simulated, it needs to be not only time-dependent, but also three-dimensional. Without three-dimensionality, the vortex stretching and tilting mechanisms, which are so important in real flows, cannot be appropriately simulated. The necessity to do 3-D calculations dramatically raises the costs of the calculations, both in terms of computer memory and in terms of run time. Turnaround time for circular cylinder calculations on eight processors of an SGI Origin 2000 can easily be multiple weeks. Grid and time-step studies therefore turn the activity into a very long process, even for such a relatively simple flow. The time can be significantly reduced through the use of more extensive parallel processing, but the parallel efficiency must be maintained. These explorations of these methods are ongoing at Langley, and therefore the most thorough testing, some aspects of the model raise questions as to its suitability for actuator-type computations. In particular, the fact that the boundary between the RANS region and the LES region is entirely dictated by grid size and distance from the wall makes the simulation of near-wall recirculation regions difficult. Other versions of composite models may work better in flows where near-wall recirculation regions are anticipated.

5.3 Langley Effort

The efforts at Langley in this area have concentrated on the development of appropriate validation experiments and also in doing some of the careful studies needed to calibrate the performance of DES with different codes. Some additional developmental work on new families of composite LES/RANS models has been funded by NASA Langley Research Center. Although the DES model has by far the greatest following, and therefore the most thorough testing, some aspects of the model raise questions as to its suitability for actuator-type computations. In particular, the fact that the boundary between the RANS region and the LES region is entirely dictated by grid size and distance from the wall makes the simulation of near-wall recirculation regions difficult. Other versions of composite models may work better in flows where near-wall recirculation regions are anticipated.

Another important role played by Langley is to act as a clearinghouse for ideas related to unsteady turbulence models. Langley researchers can provide appropriate suggestions and criticism of emerging concepts so that the developers can address these concerns early in the models' development. As an example of this interaction, the consistency concerns
discussed earlier are now being addressed in at least one new model under development.

At NASA Langley Research Center a concerted effort is underway to measure and compute the flow in active flow control devices. This effort brings together experimentalists, turbulence modelers, and computationalists who all have been working together to develop appropriate techniques for the simulation of actuator flows.

6 Actuator Boundary Conditions

6.1 Overview: Current Practices

Interest in active flow control for drag or noise reduction, flow vectoring, mixing enhancement, and separation control has stimulated the recent development of innovative synthetic jet actuators that create localized disturbances in a flowfield. Synthetic jets are generated by a dynamic fluid actuator consisting of a cavity enclosed by one (or more) moving diaphragm(s) driven into transverse oscillations at their resonance frequency. The distinctive feature of these actuators is that they have minimal power requirements and have jet-like characteristics without the need for mass injection. Although net mass injection into the overall system occurs over each cycle of operation, the momentum transfer into the embedding flow is nonzero. These features enable synthetic jets to effect significant global modifications in the embedding flow on scales that are one to two orders of magnitude larger than the characteristic length or force scale of the jets themselves.

In recent years, because of considerable improvements in computer resources, more attention is being devoted to numerical simulation and optimization of synthetic jet actuators.82, 46, 58, 33, 147, 29, 64, 89, 118, 77, 87 The presence of several temporal and spatial scales and moving boundaries in the problem makes simulation of such unsteady flows computationally very expensive. To reduce the computational cost, several approaches have been developed. All the methods can be divided into two classes: 1) simplified models without simulation of the flow inside the actuator, and 2) numerical simulation of both the exterior and cavity flows. In the simplified models, a synthetic jet generated by harmonic motion of the actuator diaphragm is modeled by using simplified boundary conditions imposed at the orifice exit. In the work of Donovan, Kral, and Cary,82 and Kral et al.,46, the flow within the cavity is not calculated and the perturbation to the flowfield is introduced through the wall-normal component of velocity at the orifice exit

\[ v(x, y = 0, t) = V_A(x)\sin(\omega t) \]
\[ u(x, y = 0, t) = 0 \]  \hspace{1cm} (21)

where \( x \) and \( y \) are the streamwise and cross-stream directions, respectively. Different spatial variations of \( V_A(x) \) over the orifice have been considered. Numerical experiments 45, 82 indicate that a "top hat" distribution most closely matches the experimental data. A modified boundary condition on the pressure at the orifice is introduced through a consideration of the normal momentum equation. Taking into account time-harmonic velocity perturbations, this condition, obtained under the assumption that the flow is incompressible, becomes

\[ \frac{\partial p}{\partial y} = -\rho V_A(x)\omega \cos(\omega t) \]  \hspace{1cm} (22)

The numerical simulations based on the boundary condition (eqs. 21 and 22) show good qualitative agreement with the experiments.122, 33, 131 A similar approach modeling the actuator as a blowing/suction boundary condition, which can be fully specified in advance of the calculation, is used by others.29, 58, 64, 89

An alternative technique for modeling synthetic jet actuators has been proposed by Carpenter et al.33 This theoretical model for the actuator is based on classic thin plate theory for the diaphragm dynamics. The flow through the orifice is modeled using unsteady pipe-flow theory. This approach is based on the assumption that streamlines in the orifice exit are parallel to its axis, which is an adequate approximation if the length-to-diameter ratio is much larger than unity. The governing equation for the orifice flow is given by

\[ \rho \frac{\partial v}{\partial t} + \rho v v_r = -\frac{\partial p}{\partial y} + \mu \frac{\partial}{\partial r} \left( \frac{1}{r} \frac{\partial v}{\partial r} \right) \]  \hspace{1cm} (23)

where \( y \) and \( r \) are the radial and axial coordinates, respectively, \( v \) is the axial velocity, \( l \) is the orifice length, and density \( \rho \) is taken to be the instantaneous mean of the cavity and external densities. The inertial term \( \rho v v_r \) is modeled approximately by the second term on the left-hand side of equation (23). The dynamics of the air in the cavity is ignored and the pressure there is calculated by means of the perfect gas law. The cavity and the external boundary layer flowfields are linked by requiring continuity of velocity and pressure at the orifice exit. In Carpenter, Lockerby and Davies,33 only the blowing phase of the actuator dynamics has been studied.
The second class of methods is based on a direct numerical simulation of the entire problem including the flow inside the actuator. Rizzetta, et al.\textsuperscript{118} solve the unsteady compressible Navier-Stokes equations in the external region, the cavity itself, and the throat of the actuator on separate grids that are linked with each other through a Chimera methodology. The membrane motion is simulated by varying the position of appropriate boundary points. As follows from the numerical calculations presented in reference [118], the internal cavity flow becomes periodic after several cycles. Therefore the velocity profile across the jet exit at each time-step was recorded and was used as a boundary condition in subsequent runs involving the external domain only. For computations that consider only the upper exterior domain, the transverse and span-wise velocity components (orthogonal to jet axis) are set to zero and the inviscid normal momentum equation, expressed as

$$\frac{\partial p}{\partial y} = -\rho \left( \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial y} \right)$$ \hspace{1cm} (24)

is used to establish the pressure. The orifice exit density is extrapolated from the interior solution. This approach provides more accurate description of the flow details at the orifice than the simplified boundary condition of equations (21) and (22). Similar direct numerical simulation of the external and cavity flows has been performed by Joslin et al.\textsuperscript{77} and shown good agreement with experimental results.

To avoid the integration of the Navier-Stokes equations on a moving grid, an alternative technique is used by Lee and Goldstein.\textsuperscript{87} The main idea of the method is to impose a localized body force along desired points in the computational mesh to bring the fluid there to a specified velocity so that the force has the same effect as a solid boundary. The desired velocity is incorporated in an iterative feedback loop to determine the appropriate force. For a moving boundary with velocity \(V_A(x,t)\), an expression for the body force is

$$F(x,t) = \alpha \int_0^t (v - V_A) dt' + \beta (v - V_A)$$ \hspace{1cm} (25)

where \(v\) is the fluid velocity, and \(\alpha\) and \(\beta\) are user-defined constants that are negative and can be treated as the gain and damping of the force field with dimensions of \(M/(L^3T^2)\) and \(M/L^3T\), respectively. This approach allows a moving diaphragm without using a time-dependent coordinate transformation.

### 6.2 Bottlenecks

In spite of the fact that the methods mentioned above have successfully been used for modeling synthetic jet actuators, several issues persist. One of the main drawbacks of the first class of methods is that the simplified boundary conditions do not provide conservation of mass, momentum, and energy through the actuator orifice. Because these methods use the normal momentum equation to calculate the pressure, whereas the other quantities are extrapolated or prescribed analytically, this boundary condition does not satisfy the governing equations at the boundary and, therefore, does not provide the conservation properties. Another disadvantage of the boundary condition [eqs. (21) and (22)] is its inability to account for changes in the pressure field caused by the interaction of the external boundary layer and actuator. Furthermore, as has been shown in Lee and Goldstein,\textsuperscript{87} the real streamwise velocity profile and the velocity component in the cross-flow direction are far from the analytical expressions of equation (21).

The main problem associated with the second class of methods is complexity/cost. The numerical calculation of the cavity flow requires a large number of grid points. For geometries fitted with multiple actuators, grid requirements for the actuators could exceed those of the exterior flow, and would contribute extensively to the computational cost. Another consideration is the actuator Mach number, which varies from 0.001 (near the diaphragm) to 0.1 (at the orifice exit). This variation of the flow parameters from fully incompressible in the actuator to fully compressible in the exterior region imposes very severe requirements on a numerical method and increases the algorithm complexity.

### 6.3 Langley Effort

As follows from the literature overview presented above, the research in the area of active flow control is of empirical nature, mainly due to the computational cost involved and lack of confidence in computational methods for such complex time-dependent flows. A strong effort\textsuperscript{147} is currently underway at Langley toward constructing a new methodology that combines the accuracy and conservation properties of the simulation methods with the efficiency of the techniques based on simplified boundary conditions. In contrast to the methods found in the literature, the new approach uses a reduced-order model of 2-D or 3-D actuators. In other words, the multidimensional actuator is simulated by solving the time-dependent 1-D Euler equations similar to those used for the quasi-one-dimensional nozzle.
problem. The simplified actuator model has several advantages. First, this approach provides conservation of not only mass, but also momentum and energy. Second, the new method is much more efficient in terms of computational time compared with the 2-D or 3-D numerical simulation of the flowfield in the actuator. Third, this reduced-order model retains some multidimensional features of the realistic actuator that are governed by the area variation of the quasi-one-dimensional nozzle, and, therefore, can be used for qualitative study of the resonance characteristics of the actuator. As follows from the comparison presented in figure (3), the conventional boundary conditions based on the normal momentum equation do not provide mass conservation. The maximum mass rate error, which occurs during the suction stage, is of the order of 15 percent if the normal momentum equation is used as a boundary condition for pressure. In contrast to the conventional approach, the new method provides conservation of mass, momentum, and energy. As a result, the mass rate error is reduced by one to two orders of magnitude compared with that obtained with the blowing/suction boundary condition [eqs. (21) and (22)]. These preliminary results are very encouraging. In our future work, we will focus on calibration of the new methodology with readily available experimental data and numerical simulation of time-dependent flows encountered in active flow control environment.

7 Conclusions

The current status of time-dependent algorithms is presented. Special attention is given to algorithms used to predict fluid actuator flowfields. The overview begins by considering algorithmic issues that could greatly improve the temporal efficiency of actuator simulations. The general state of time-dependent turbulent models for nonstationary flows is then assessed. Finally, an efficient new fluid actuator boundary condition is presented. Each section begins by describing the current state of the art including notable impediments in the field, and concludes with a summary of current Langley efforts.

Profound improvements in the efficiency of temporal algorithms could be achieved in the next decade. Notable leverage in time-dependent methods exists in the following algorithmic areas: 1) implementation of high-order \( p \geq 3 \) temporal schemes, 2) implementation of high-order \( p \geq 3 \) spatial algorithms, and 3) convergence acceleration techniques for complex high-Reynolds number flows. Significant impediments exist in each of these three categories.

High-order schemes need huge time steps to utilize their full potential. Algebraic solvers with rapid, time-step independent, convergence rates are necessary for these schemes. The principal impediment facing the implementation of high-order temporal schemes is the need for robust and rapid algebraic equation solvers. Error estimation, variable time-stepping and automated iteration termination will immediately follow. Current fourth-order schemes are asymptotically close to being optimal, and further increases in efficiency are difficult to obtain.

High-order spatial operators must be flexible enough to accommodate complex geometries, grid adaptation and nonlinear instability. Methods utilizing compact locality (unstructured methods) are advantageous when addressing complex geometries and grid adaptation. Unstructured high-order formulations include Finite-Elements (FE), \( k \)-exact Upwind Finite-Volume (FV), or \( k \)-exact ENO FV. The principal impediment facing the implementation of high-order spatial operators is their lack of nonlinear stability resulting from unresolved features and discontinuities. Extensive research is currently underway within the FE community. Some formulations possess a nonlinear \( L_2 \) stability property. It is still unknown whether this stability is a strong enough for general purpose solvers, or whether stronger stability conditions (TVD, TVB, ENO) must be pursued.
A wide variety of different iterative methods are currently used. Nevertheless, the convergence rate of current state of the art iterative solvers is poor. Complex high Reynolds number 3-D simulations converge at per cycle rates in the range 0.96 – 0.98. The principal impediments for rapid convergence are high aspect ratio cells in the turbulent boundary layers, and low Mach number regions in the flow.

Conventional URANS turbulence models are not very accurate for nonstationary turbulent flows. A new class of composite LES/RANS schemes has been developed to address these inadequacies. Among these is the detached eddy simulation (DES) approach, and is currently being verified and validated. The DES scheme is very expensive because it requires a full 3-D time-dependent simulation of the flow in question. The principal impediment facing all composite approaches is validation against high quality experimental data. The accuracy and consistency of the blended region between the outer LES and inner RANS regions is a concern with composite methods. The accuracy of this region is extremely important in fluid actuator simulations.

An optimistic estimate indicates that successful research could improve overall efficiency by O(10^{1/2}) for temporal algorithms, by O(10^{3/2}) for spatial algorithms, and by O(10^{1/2}) for convergence acceleration. Improvements in the modeling of fluid actuators could account for a factor of two. Moore’s law predicts that increases in computer hardware will yield O(10^4) improvements. All these effects can be combined in a multiplicative sense to yield potential improvements of O(10^6). Improvements of this magnitude would allow us to do 3-D optimization studies based on DES turbulence models, including fluid actuator design and resonance, and actuator placement and coupling studies.

References


61. P. N. Brown, B. D. Byrne and A. C. Hindmarsh, VODE: A Variable Coefficient ODE


