ADVANCES IN COMPUTATIONAL CAPABILITIES FOR HYPersonic FLOWS

Ajay Kumar, Peter A. Gnoatto, James N. Moss
and
J. Philip Drummond

Aero- and Gas-Dynamics Division
NASA Langley Research Center
Hampton, VA 23681
U.S.A.

Presented at the AGARD Conference
on
Future Aerospace Technology in Service to the Alliance

Palaiseau (near Paris), France
April 14–17, 1997
ADVANCES IN COMPUTATIONAL CAPABILITIES FOR HYPERSONIC FLOWS

Ajay Kumar
Peter A. Gnofo
James N. Moss
J. Philip Drummond
Aero- and Gas-Dynamics Division
NASA Langley Research Center
Hampton, VA 23681
U.S.A.

1. SUMMARY
The paper reviews the growth and advances in computational capabilities for hypersonic applications over the period from the mid-1980's to the present day. The current status of the code development issues such as surface and field grid generation, algorithms, physical and chemical modeling, and validation is provided. A brief description of some of the major codes being used at NASA Langley Research Center for hypersonic continuum and rarefied flows is provided, along with their capabilities and deficiencies. A number of application examples are presented, and future areas of research to enhance accuracy, reliability, efficiency, and robustness of computational codes are discussed.

2. INTRODUCTION
Over the years, worldwide research and technology development in hypersonics seems to follow a cyclic pattern with activities peaking around specific missions and technology programs such as Apollo, Space Shuttle, and more recently, NASP and HERMES. The current hypersonic activities are predominantly focused around the development of low cost reusable launch vehicles and associated propulsion systems. This new generation of vehicles is primarily driven by economic considerations, although it does have military implications. The cyclic pattern of technology development in hypersonics has had severe impact on the scientific community in its ability to maintain necessary skills and test facilities during the period when there are no approved hypersonic projects. It has been even more difficult to recreate them when new projects do appear. However, since the early 1970's, a new engineering tool, known as Computational Fluid Dynamics (CFD), started to emerge for aerospace applications, and its growth through the mid 1980's was not tied to any specific technology program or mission. Significant growth in CFD capabilities took place during this period, ranging from the solution of inviscid incompressible flow fields to the solution of Navier-Stokes equations for high Reynolds number and Mach number [1,2,3]. Even with this rapid growth in CFD, it did not yet become a routine analysis and design tool for aerospace applications. However, in the mid 1980's, CFD was accepted as an attractive analysis and design tool mostly due to demands of the programs such as the National Aero-Space Plane (NASP) Program and aerobraking research program, and not necessarily because it had become a mature design tool. These programs had no other design tools immediately available to them. Due to lack of available ground facilities for testing at high Mach numbers (the ones that were available did not simulate the flight enthalpies at these high Mach numbers) and due to the complex and integrated nature of the flow field, traditional wind-tunnel based design techniques were not adequate. CFD was offered as a promising alternative for analysis, design, and even optimization of hypersonic vehicles and propulsion systems. However, up to around this time, most of the available CFD capabilities were either two-dimensional Navier-Stokes codes with limited physical/chemical modeling or were simplified codes (viscous-shock layer, parabolized Navier-Stokes) with limited applicability. Fully three-dimensional Navier-Stokes codes were just beginning to appear, and they were generally applicable only for perfect gas flows. In addition, none of the codes had been properly validated.

However, in the period between 1985 and 1990, there was a significant increase in the application of CFD for modeling hypersonic flows. Industry slowly accepted it as a viable alternative, and significant resources were invested in the development of CFD codes by both industry and government labs. Several fully three-dimensional Navier-Stokes codes, along with their space marching versions, were developed with detailed chemistry modeling, both for high-speed combustion and aerothermodynamic applications. Most of these codes used algorithms developed in the seventies and early eighties, namely, implicit or explicit central differences methods (including multistage methods) and upwind methods. Simple phenomenological turbulence models were employed, none of which was truly applicable to either nonreacting or reacting hypersonic flows. The Direct Simulation Monte Carlo (DSMC) based methods for computing low density, high altitude flows also experienced a much expanded role in aerospace applications including complex three-dimensional simulations.

From 1991 to the present day, only modest progress has been made on either the algorithms or the physical/chemical modeling used in hypersonic CFD codes. With the exception of modeling in DSMC methods, most codes continue to use the same algorithms and models with minor refinements, and little data is available to validate these refinements. However, significant progress has been made on surface and flow field discretization using structured and unstructured grid methods, and significant gains in efficiency are being achieved through the increasing use of parallel computers and "smart" ways of applying available CFD codes to a given problem.

The paper reviews the growth and advances in CFD for hypersonic applications over the period from 1985 to the present time. Since the flow on a hypersonic vehicle may include domains that vary all the way from subsonic, high pressure and temperature stagnation regions behind strong bow shock waves to high speed, low density (high Knudsen number) regions, the discussion of computational capabilities has been divided into the following two main areas:

Hypersonic continuum flow capabilities: Most of the discussion in this area is focused on the solution of Reynolds-
Averaged Navier-Stokes (RANS) equations with appropriate modeling of various physical and chemical phenomena present in the flow field. The applications include aerothermal dynamic analysis of the external flow over hypersonic vehicles and planetary entry probes and internal flow analysis of the hypersonic airbreathing propulsion flow path.

Rarefied and free-molecular flow capabilities: The capabilities are required for computing hypersonic flow over a wide spectrum of conditions ranging from low density flow at high altitude to relatively high density flow with small characteristic dimensions. Under these conditions, the continuum modeling of the flow becomes deficient, and it is necessary to recognize the molecular structure of the gas. A number of codes have been developed to analyze such flows using the DSMC method and are being used extensively to analyze flow over hypersonic vehicles, planetary entry probes, etc.

In the following sections, basic issues in the development and application of each of the preceding computational capabilities are discussed. A brief summary of some of the currently available popular CFD codes is provided in each of the preceding areas along with their capabilities and deficiencies, computational speed, and example applications. Future areas of research to enhance accuracy, robustness, and reliability of the computational codes are also discussed.

3. HYPERSONIC CONTINUUM FLOW CAPABILITIES

The growth of RANS-based computational capabilities in the United States for external and internal, continuum hypersonic flow during the period of 1985 to 1990 was driven primarily by generic studies of aeromanuevering and aeroassist ed orbital transfer vehicles, Space Transportation System (STS) support, defense interceptors design, and the NASP program. In Europe, programs such as HERMES, Sanger, and HOTOL provided similar impetus. These computational capabilities were further developed to maturity during the period 1991 to the present, and their applications during this period include closeouts or continuations of several of the preceding programs as well as some new programs such as X-33 (a technology demonstrator for a Reusable Launch Vehicle (RLV)), X-34 (a Mach 8 test vehicle), X-38 (an assured crew return concept from the International Space Station), Hyper-X (a subscale vehicle to flight demonstrate integrated scramjet engine performance at Mach 5, 7, and 10), and a number of planetary entry missions (Mars Pathfinder, Mars Penetrator, Stardust, etc.).

In general, hypersonic flows, with the presence of strong shocks and complex interactions and the need to model complex physical and chemical phenomena, present significantly larger challenges in the development of computational capabilities than the subsonic or supersonic flow. However, even within hypersonic vehicles, design of a hypersonic airbreathing aircraft places a different set of demands on computational tools than does a rocket-powered vehicle. These demands stem largely from the following dominating requirements for the hypersonic airbreathing aircraft:

- Close integration of the airframe and propulsion system driven by very small performance margins, which is the difference of two large numbers, thrust and drag, of almost equal magnitude. The actual value of this performance margin is at least an order of magnitude smaller than either the thrust or drag, thus necessitating highly accurate prediction of vehicle aerodynamics and engine performance. Also, the close integration of airframe and propulsion systems produces highly interacting, complex three-dimensional viscous external and internal flow, thus further complicating the prediction of such flows.
- Integrated vehicle aerodynamics and engine performance optimization over a Mach number range.
- Vehicle reusability with minimum refurbishment between flights.

It is obvious from the preceding requirements that a nose-to-tail computational capability for a highly integrated hypersonic airbreathing aircraft has to deal with flow fields of substantially greater complexity, both geometrically and physical/chemical modeling wise, than that for a rocket-powered vehicle (although physical and chemical modeling requirements for high energy entries (≥8 km/sec.) of aerocapture and aeromanuevering vehicles may involve an even higher degree of complexity). The use of CFD as a tool in the analysis and design of such vehicles requires development of advanced codes which can deal with preceding requirements of complexity of the flow field and geometry, and are accurate, efficient, and robust. In order to meet these requirements, the basic issues that need to be addressed in the development of a CFD code are:

- Physical and Chemical Modeling
- Surface Modeling and Grid Generation
- Computational Algorithms
- Code Validation

Once a code is developed and validated, proper strategy must be applied for its efficient application in the analysis, design, and optimization processes. In what follows now, a brief discussion is given on each of these issues for the RANS level of modeling.

3.1 Physical and Chemical Modeling

In order to address all the modeling issues present in the flow field of a hypersonic vehicle, discussion is focused on a vehicle propelled by an airbreathing engine. For this discussion, the vehicle is divided into four components, namely, forebody, inlet/isolator, combustor, and nozzle/afterbody. Most modeling issues for a rocket-propelled vehicle or a planetary entry probe are discussed in the forebody part of this section.

3.1.1 Forebody

The Mach number on the forebody of a hypersonic vehicle varies from low subsonic to hypersonic and the pressure and temperature also vary over a wide range. The flow features present in the forebody flow field include the bow shock, embedded shocks due to the presence of compression ramps and other surfaces, laminar-transitional-turbulent boundary layer, and shock-boundary layer interaction and associated flow separation. It is essential to accurately predict the shape of the bow shock, all embedded shocks, and the entire shock layer profile on the forebody, including the viscous layer, so that forebody wall properties such as pressure, skin friction, and heat transfer, as well as the mass and momentum flux entering the engine inlet, can be calculated. The forebody wall properties are required for accurate prediction of forebody aero and heating loads, whereas the entire engine performance depends upon the
accuracy of mass and momentum flux predictions at the face of the inlet. The reusability of the vehicle implies a nonablative surface, thus simplifying the wall boundary conditions which need not to account for surface ablation and associated shape change. However, sophisticated wall boundary conditions, which include surface catalysis and possible thermal deformation of the forebody surface, may also become important at higher Mach numbers.

It is desirable to maintain a laminar boundary layer on the forebody as far back as possible before it transitions to fully turbulent to reduce drag and heating loads. This requires accurate prediction methods for transition onset that can be used in designing the forebody shape. Most currently available transition onset prediction methods use empirical methods, which are based on the compressible linear stability theory [4,5]. These methods require a transition onset database from quiet wind tunnels and flight tests to correlate the value of N. Since there is very limited flight test data available and since most wind tunnel data is taken in noisy environments and at non-representative flight conditions, the transition onset prediction methods are not reliable, yet they are being used in the preliminary design of hypersonic vehicles. Once the transition onset takes place in the compressible boundary layer, there is usually a transition region before the boundary layer becomes fully turbulent. At hypersonic Mach numbers, this transition region is normally quite long and, therefore, has significant impact on the total drag and heating loads on the forebody. This region may also contain areas of heating rates higher than those encountered in the fully turbulent region. The extent of the transition for hypersonic boundary layers is in the poorest state, as compared to all other type of modeling. Most computational codes today either model this transition region in some ad hoc manner or use models that were developed for incompressible flow. The turbulent boundary layer is modeled mostly by either an algebraic or a two-equation eddy viscosity model. Some type of empirical compressibility correction is applied to these models; however, little or no consideration is given to the presence of real gas effects, shock-boundary layer interactions, surface catalysis effects, etc. In general, accurate modeling of transition and turbulent boundary layers needs the most attention if we have to increase the accuracy and reliability of the computational predictions.

In the preceding discussion, the presence of real gas effects on the forebody has been mentioned, which normally become important as the flight Mach number increases in the range of 10 and above. These effects, in their simplest form, are sometimes approximated by an effective gas that yields the correct density ratio across the shock. This approach may provide useful results for aerodynamic analysis if the perfect gas option is the only one available in the code; however, it is not satisfactory for evaluation of aerothermodynamics. High temperature effects are more appropriately handled with chemical equilibrium (CE) or chemical nonequilibrium (CNE) models. In the CE models, the functional dependence of pressure on density and energy for a generalized gas mixture, in which the elemental mass fractions are known constants or are solved from the elemental continuity equations, can be determined by the method of free energy minimization or equilibrium constants. In the CNE models, the species concentrations are determined by solving species continuity. Mixture pressure is defined as a summation of the partial pressures of each constituent. The partial pressure of free electrons is computed using the electron temperature if different from the heavy particle temperature.

Within CNE models, the chemical source terms may be expressed as functions of thermal equilibrium (TE) or thermal nonequilibrium (TNE) chemical kinetic models [6,7,8]. The TNE models may be divided into four subclasses. The simplest one is the two-temperature model in which it is assumed that the translational and rotational energy content of all heavy particles is defined by temperature T. The vibrational energy content of all molecules is defined by temperature T_v. The electronic and free electron translational energy modes are usually assumed to be in equilibrium with the vibrational modes at temperature T_v. The next logical TNE subclass is the three-temperature model. It is identical to the two-temperature model except that the electronic and free-electron translational energy partition is defined by a third temperature T_e. This model has been considered primarily for high-speed earth entry characteristics of return from geosynchronous earth orbit or above. The third subclass of TNE is the multi-temperature model which is again similar to the two-temperature model except that each molecular species retains its own vibrational temperature T_v. Such models are most often used in the code validation process using ground-based experimental data in which separate vibrational temperature measurements have been made.

All of the above subclasses of TNE models invoke an assumption of equilibrium partitioning of energy within two or more groupings of internal modes. This assumption can be relaxed further by discretizing each species into multiple sub-species (excited states) as a function of vibrational and/or electronic quantum number(s) [9]. In this fourth subclass of TNE models, a unique translational temperature is still appropriate within the context of a continuum flow simulation; however, the role played by modal temperature in defining energy distribution within the mode is replaced by a discrete accounting of various excited states. Each state requires its own continuity equation, with appropriate models for transport and source terms.

To solve the governing equations for the diffusion of mass, momentum, and energy in a gas mixture, thermodynamic and transport properties are required. The specific heat for each species may be defined, for example, by a fourth-order polynomial in temperature [10]. Transport properties of a gas mixture are defined as functions of constituent species transport properties and respective mole fractions in the mixture [11]. Individual species transport properties are defined either by Sutherland’s law or as functions of collision cross sections [12]. Direct curve fits are also available for air viscosity and conductivity under thermochemical equilibrium [13]. The Chapman and Cowling law is used to determine the binary diffusion coefficients which describe the diffusion of each species into the remaining species [14]. Knowing the diffusion coefficients, the diffusion velocity of each species is determined by solving the multicomponent diffusion equation [15]. However, in most engineering calculations, it is assumed that each species present in the flow has the same diffusivity, and pressure and thermal diffusivity may be ignored. Under this assumption, the diffusion of each species into the remaining species varies only with its concentration gradient and can be described by Fick’s law.
3.1.2 Inlet/Isolator

For hypersonic airbreathing vehicles, the inlet/isolator of a scramjet supplies the combustor with air at a specified pressure, velocity, and flow uniformity. The flow physics in this region include three-dimensional compressible turbulent boundary layers, transitioning boundary layer on some inlet surfaces, oblique shocks, shock-boundary layer interactions and associated flow separation, localized high leading edge thermal loads due to shock-shock interactions, unsteady flow due to possible inlet unstart and module to module interactions, and internal-external flow interactions due to the spillage as a result of the requirement that the inlet need to operate over a Mach number range. In addition, modeling of flow in this region may be further complicated by the presence of wings. Even though in most preliminary inlet design cycles, inviscid Euler codes iterated with boundary layer codes may be used, the physics of the inlet flow field requires the use of full Navier-Stokes codes with compressible turbulence models. In most current codes for inlet flow field analysis, turbulence is modeled by either algebraic or two-equation turbulence models with empirical compressibility corrections and wall functions. The flow is assumed to be fully turbulent on all surfaces of the inlet. Very limited computational studies have been conducted of inlet unstart and module to module interactions [3] due to the unsteady nature of the flow, inability of the codes to model strong shock-boundary layer interactions and large regions of separated flow, and large computational resources required to conduct such calculations. As for the forebody, the most limiting area for accurate modeling of inlet flow fields is the development of advanced transition and turbulence models. Promising work is currently underway to develop new algebraic Reynolds stress turbulence models which have the computational efficiency of a two-equation turbulence model but include more physics.

3.1.3 Combustor

The flow field in the combustor of a scramjet engine is probably the most complex amongst all the components of the hypersonic vehicle. It is characterized by much of the flow physics of the inlet/isolator region, but it is further complicated by a wide range of flow velocities inhomogeneously distributed throughout the combustor, small and large scale vortical flows for fuel-air mixing, finite rate chemical reactions for ignition and combustion resulting in high temperatures and heat fluxes, anisotropic and nonequilibrium transfer of turbulence energy, and turbulence/chemical kinetics interactions. Computations of combustor flow fields require the use of RANS based codes; however, a parabolized Navier-Stokes code may be used in some analyses. Turbulence is again modeled using algebraic or two-equation models with empirical compressibility corrections and wall functions. Few codes account for turbulence chemistry interactions through an assumed probability density function approach. Chemical reaction is modeled with reduced reaction set finite rate models. For the hydrogen-air reactions occurring in a hydrogen-fired scramjet, a typical reaction mechanism includes nine chemical species and eighteen chemical reactions, although other less or more elaborate mechanisms may be used depending upon the case [16]. In a hydrocarbon-fired scramjet, combustion mechanisms are much more complex and require carefully reduced mechanisms to allow practical computations.

In the analysis of a combustor, apart from predicting the overall flow field, there are a number of locations where highly detailed analysis of very localized processes is required such as in the neighborhood of fuel injectors and flame holders. Accurate and detailed prediction of such localized flow regions is required if computational tools are to be used in the design of an efficient combustor.

3.1.4 Nozzle/Afterbody

The flow field in the nozzle/afterbody of a hypersonic vehicle retains some of the physics of the inlet and combustor, but also includes additional requirements of modeling highly expanding nonuniform flow that may reestablish and then retransition back to turbulent flow, significant divergence and skin friction losses, energy-bound chemical radicals that will not relax in the nozzle, and excited vibrational states and their relaxation. Nozzle/afterbody flow field calculations are usually made with either an Euler code iterated with a boundary layer code or a parabolized Navier-Stokes code. Calculations may also be continued with the same Navier-Stokes code that has been used in the analysis of combustor flow field. Finite-rate reaction chemistry is still required in the nozzle to assess the degree of reaction that continues to take place and to determine the extent of recombination reactions that add to the available thrust. The reduced kinetics models currently being applied to nozzle/afterbody flow appear to be reasonably accurate, although some further work to improve the description of recombination may be warranted. At present, there is little or no modeling of the reinitialization and retransition of the boundary layer that might occur in the region of large expansion. Overall, nozzle/afterbody flow field computations are much less demanding than either the inlet/isolator or combustor flow field.

3.2 Surface Modeling and Grid Generation

With the acceptance of RANS technology in analysis and design of complex configurations, the requirements for rapid solutions dictated that all elements of the solution process be refined and streamlined. Surface modeling and grid generation appear to be the most significant contributors to the total time required to generate a solution. Reference [17] summarizes the advances and remaining challenges in this area. A surface modeling data-exchange standard has been developed to provide a link to computer-aided-design surface modeling packages, and the surface model definition is now routinely incorporated into grid generation packages to maintain surface fidelity as the grid is enriched or adapted. An extensive array of block-structured methods, unstructured grid methods, and hybrid schemes are now available, although no single method has emerged as the preferred approach. The block structured methods, including patched and overset grids, are the most efficient methods for high Reynolds number simulations, but are the most labor intensive in terms of grid generation; automated methods are being developed but are still in their infancy. However, to speed up the grid generation process, libraries of efficient topologies for classes of configurations are being developed which can be used to automatically generate a new three-dimensional grid for changes to an existing model in a very short time.

The unstructured grid methods have emerged as the methods of choice for nonlinear inviscid simulations because of their relative speed and lack of the need of user interaction in the grid generation process [18]. The viscous unstructured grid methods are also becoming available for complex three-dimensional configurations [19]. However, for high Reynolds
number viscous flow, a hybrid approach may be more useful. In this hybrid approach, a structured layer of points is developed near the surface and these near-field layers are then connected to the outer field through an isotropic distribution of unstructured grid. Reference [20] describes this advancing-layer, advancing-front hybrid grid generation method and uses it successfully for a three-dimensional viscous calculation.

3.3 Computational Algorithms
Some of the most important algorithmic advances for the computation of hypersonic flows have been in the development of upwind and non-oscillatory schemes for improved shock capturing. Central difference schemes with upwind-biased or non-oscillatory dissipation operators are included in this class of algorithms. Upwind schemes evolved more to satisfy an intuitive valuation that consistency between physics and numerics is important. High-resolution, non-oscillatory schemes more directly confront issues of accuracy in the region of high gradients such as in the vicinity of shocks and shear layers. Prior to these advances, flows with strong shocks (pressure ratio in the range of 100 or more) could not be computed accurately with shock-capturing methods based on central differencing; either the requisite dissipation excessively smeared the shock or Gibb's phenomena caused large enough oscillations to generate negative temperatures in the neighborhood of shocks. Shock-fitting methods have advanced to a much lessor extent as compared to more flexible, easily coded shock-capturing methods.

In general, the RANS equations are solved to steady state by properly posed, time dependent problem and marching the solution to large time with steady state boundary conditions. In hypersonic flows, this approach has its principal advantage in being able to maintain fully conservative differencing, thus providing accurate shock capturing. The algorithms that are in use for time-stepping can be classified as either explicit or implicit schemes. Although explicit schemes are relatively easy to code, they are prohibitively expensive for the highly stretched grids associated with high Reynolds number viscous simulations because the explicit time step scales as the square of the mesh size for pure diffusion model problems. However, one of the most popular and widely used algorithms through the early to mid eighties was the explicit MacCormack algorithm [21]. An implicit algorithm, such as the Beam and Warming [22] algorithm, allows to advance the solution with much larger time step as compared to an explicit algorithm and also has the flexibility in the selection of the actual time step. This aspect becomes particularly important as the predominantly reduced frequencies become small.

The Runge-Kutta time-stepping method, originally applied to Euler equations by Jameson, Schmidt, and Turkel [23] in 1981, is another extensively used method for steady and unsteady viscous flow computations. Although the basic algorithm is explicit, residual smoothing is used to extend its stability limit. The three-dimensional RANS technology has become even more acceptable by the aerospace industry with the advancements in multigrid algorithms [24] which have significantly accelerated the convergence to steady state over a single-grid algorithm. This technology is still not fully developed, especially for hypersonic flows due to difficulties in the treatment of chemical source terms and strong shocks, but the prospects exist for considerable further enhancements to the convergence rate [25,26]. Local preconditioning [27] is another algorithmic enhancement that is currently under development to address problems associated with convergence and truncation errors in very low velocity flows, such as in the stagnation region of a blunt body in hypersonic flow or in the regions with significant viscous effects. Here again, much work remains to establish optimum performance of local preconditioning.

Most central-difference algorithms require explicit addition of dissipation terms, such as the second- and fourth-difference operators of Jameson et al. [23], which provide a third-order dissipation in smooth regions of the flow and a first-order dissipation in discontinuous regions of the flow. A pressure gradient-based switch is used to turn off the fourth-difference dissipation in discontinuous regions to avoid small scale oscillations. A modified pressure switch was developed by Turkel et al. [25] and Vatsa et al. [26] for hypersonic flows to avoid negative density or pressure near strong shocks. An alternate way to add dissipation is through the use of flux-difference or flux-vector splitting, upwind-difference methods which have more physical, inherent dissipation and do not require it to be added explicitly [28].

As discussed earlier, hypersonic flows require modeling of real gas, equilibrium or nonequilibrium chemical phenomena which involve chemical energy source terms and additional species equations. In cases where the chemical time scale is much longer than the fluid dynamic time scale, the species equations can be solved in a weakly coupled fashion as a subset of the Navier-Stokes equations. Even though this approach is computationally efficient, in most cases the chemical time scale is of the same order or smaller than the fluid dynamic time scale. In these situations, the species equations are solved in a fully coupled fashion with the remaining equations. The species equations have a chemical source term that describes the production and loss of each chemical species in the reaction process. This source term normally causes the overall system of governing equations to be numerically stiff, when the chemical time scale is relatively smaller than the fluid dynamic time scale, and is typically solved implicitly with either explicit or implicit time-stepping methods. With implicit discretization of the source term, the resulting system is rescaled by a matrix containing the source Jacobian that allows the equation system to be advanced in time at nearly the fluid dynamic time scale [29,30].

3.4 Code Validation
Before a computational code can be used in the development of aerospace vehicles, it is necessary to determine its accuracy in predicting the flow field through a validation process. This validation process is required to ensure that the physical modeling used in the code approximates the actual flow physics and that the algorithm employed in the code to solve the governing equations is not adding significant numerical error. In order to validate a code and establish its range of applicability, it is necessary to compare numerical results from the code with some well-posed experiments. Since computational codes are being applied to highly complex and interacting flow field analyses, it is no longer adequate to compare against global experimental data such as surface or other integral quantities. As an example, it was mentioned in the earlier section that it was not sufficient to just calculate the surface pressure and heating on the forebody, but was also necessary to accurately predict the whole shock layer profile entering the inlet/isolator region because the accuracy of
performance prediction of the inlet/isolator region depends upon the inflow profile.

In general, the validation process for a newly developed code first requires some internal checks for consistency to ensure that the code preserves mass, momentum, and energy and also preserves the freestream on an arbitrary grid in the absence of a body. The newly developed code can also be checked against some other established codes. However, the major step in validation of a code consists of comparison against detailed experimental data. Normally, two types of experiments are necessary to determine the accuracy of the code. The first type of experiments, called the building block experiments, are necessary to validate physical and chemical modeling. These experiments should provide measurements which not only show the deficiencies in the modeling but also help identify the aspects of models which need improvements. The second type of experiments, called the benchmark experiments, are used to determine prediction capability of the code and to establish limits and range of applicability.

As an aerospace vehicle goes through its flight envelope, the physical and chemical processes that need to be modeled change, the characteristics of the algorithm change, and the grid requirements for resolving the flow change. It is, therefore, necessary to validate the code over small ranges of the flight envelope, and its use should be limited to the appropriate range of validation to avoid large uncertainty in predictions. It must be realized that due to a lack of high-enthalpy facilities that can simulate the flight conditions, a very limited amount of experimental databases are available for hypersonic flow. This is even more true for the combustor flow field where extensive velocity, pressure, temperature, and species concentration measurements are required along with the correlation of these quantities with each other if validation of advanced turbulence and turbulence-chemistry interaction models is required. Another factor that is pushing the accuracy requirements on the measurements and, thus, tighter limits on validation, is the necessity of predicting aerodynamics and engine performance very accurately due to overall small performance margins for a hypersonic airbreathing vehicle, as discussed earlier.

Even though over the last decade or so there appears to be a general consensus on the lack of hypersonic experimental database for code validation, very limited resources have been allocated to remedy the situation. The cost of obtaining code validation quality data on a flight experiment/technology demonstrator also remains a serious obstacle. There is always a challenge to show that the cost and risk of validation data acquisition on a flight project is justified by the risk reduction in future applications.

3.5 Major Continuum Flow Codes

This section describes some of the RANS-based codes that are representative of capabilities for hypersonic flow analysis. It is not an all-inclusive list; rather, it is restricted to those codes that are currently being employed by various groups at the NASA Langley Research Center for external and internal hypersonic flow analysis.

3.5.1 LAURA (The Langley Aerothermodynamic Upwind Relaxation Algorithm)

The LAURA code was primarily developed for external aerothermodynamic analysis of flow over hypersonic vehicles and planetary entry bodies. It can solve Euler, thin-layer Navier-Stokes, or Navier-Stokes equations. Inviscid flux definition in LAURA employs Roe’s averaging [31] and Yee’s Symmetric Total Variation Diminishing (STVD) [32] for second-order accuracy away from discontinuities. Harten’s entropy fix (eigenvalue limiter) [33] is applied across cell faces. Special variations of the limiter are employed across viscous dominated boundary and shear layers. These treatments overcome problems often encountered with the baseline Roe’s method regarding the ”carbuncle” phenomenon or the baseline Harten’s method in which numerical dissipation (proportional to an unnaturally large eigenvalue) competes with physical dissipation. Central differences are used to define viscous flux. Point-implicit relaxation of the steady form of the conservation laws at each control volume in a computational plane is implemented, sweeping from plane to plane in a block and from block to block across the entire domain of interest.

The basic features of LAURA include options for five thermochemical kinetic models for 11 species, two equilibrium air models [35,36], two temperature thermal model, two algebraic turbulence models, six models for wall catalysis, a radiative equilibrium wall, discretization on up to six, simply connected, structured blocks, and mesh sequencing. Solution Jacobians may be stored out-of-core for significant reduction in memory requirement. New features offered in LAURA include options for Martian atmospheric chemistry, 1000 structured blocks with integer stride connectivity (175 block solution tested), and post-processing files generated to support integral-boundary-layer analyses of variations in surface catalysis and emissivity of the thermal protection system. A PVM/IMPI version of the code is being tested which currently supports all of the options in LAURA and most of the options in LAURA.

The LAURA code exploits macrotasking (parallel execution of a code on a shared memory machine) on Cray computers at the subroutine level. The use of macrotasking, along with provisions for asynchronous relaxation, enables exceptionally high average concurrency of tasks for LAURA. An asynchronous relaxation also enables physically motivated load balancing in which CPU cycles are concentrated in regions known to converge slowly such as separated flow regions, near wake regions, etc. An additional feature of LAURA is its built-in grid adaptation routine that simultaneously aligns the outer boundary of the computational domain with the captured bow shock and encodes near wall grid resolution required for aerothermodynamics analysis. This feature greatly simplifies the grid generation process for multiple cases over a wide range of Mach numbers, Reynolds numbers, and angles of attack on a single configuration.

LAURA has been validated with flight data from the shuttle orbiter [37,38] and with ground-based data from the Aeroassist Flight Experiment [39] and the hypersonic compression corner [40].

3.5.2 GASP (The General Aerodynamic Simulation Program)

The development of the GASP code started around 1987 under the NASP Program, and was motivated by the need for a 'nose-to-tail' analysis capability for an airbreathing hypersonic aircraft. It supports a rich variety of options for steady and unsteady solution of Euler, parabolized Navier-Stokes, thin-layer Navier-Stokes, and Navier-Stokes equations. These
options include mesh sequencing, preconditioning, approximate factorization, Line Gauss Seidel, Generalized Minimal Residual (GMRES) [41], mesh sequencing and multigrid. Inviscid flux definition in GASP employs several options, including Roe’s and Van Leer’s upwind biased formulations and central differencing with artificial viscosity. Central differences are used to define viscous flux.

Both algebraic and two-equation turbulence models with wall function options are supported. Generalized zonal-boundary interpolation is supported across zonal intersections defined by a single logical boundary. Parallel processing is supported on shared memory computer architectures. A comprehensive set of thermochemical kinetic models is offered for air chemistry, hydrogen-air combustion, and hydrocarbons in a database containing 455 reactions and 34 species. Thermal nonequilibrium may be modeled using a separate vibrational temperature for each molecule or a lumped vibrational temperature common to all molecules. The comprehensive GASP V3 Users Manual [42] and the Graphical User Interface (GUI) for problem setup and data manipulation make GASP more user friendly.

Due to the comprehensive set of physical and chemical modeling options available in the GASP code, it is being used in the analysis of the complete hypersonic vehicle flow field. It has been validated for a number of external and internal flow fields such as discussed in References [43,44,45].

3.5.3 LARCK (Langley Algorithm for Research in Chemical Kinetics)
Several RANS-based codes have been developed and used at NASA Langley for modeling the internal flow field in scramjet inlets, combustors, and nozzles. These include some earlier codes such as NASCRIN/SCRAMIN [46,47] for inlets and SPARK series of codes [48,49,50] for combustors which were developed in the early to mid-eighties. However, at present, there are two more recently developed codes which are being used extensively. One of them is the previously discussed GASP code and the other code is the LARCK code. Development of the LARCK combustor code began in 1992 to incorporate some of the algorithmic advances and geometric generalizations that had appeared in the literature during the late 1980’s and early 1990’s. Its development was driven by the need to consider more complex internal engine geometries with algorithms that were very efficient for solving high-speed reacting flows. LARCK is a multiblock, multigrid code that uses a cell-centered, second-order finite-volume integration scheme. Each region in the code is constructed of a number of blocks. Blocks within a region can be connected either elliptically or hyperbolically, whereas regions are connected parabolically. Convection terms in the governing partial differential equations are discretized in a second-order manner using either central differencing with scalar matrix artificial dissipation, or with an upwind MUSCL scheme with Roe’s approximate Riemann solver. Diffusion terms are discretized in a second-order manner with either a full gradient evaluation using Green’s theorem or an approximate gradient evaluation using finite differences.

Once the spatial terms have been discretized, LARCK has several options for advancing the governing equations in space and time. For elliptic flows, either steady or unsteady solution procedures can be used. The equations are solved using either a Runge-Kutta scheme with or without residual smoothing, a diagonalized approximate factorization scheme, or a FMG/FAS multigrid acceleration scheme using either of the above two approaches. The governing equations can also be marched in space using the pseudo time interactive method of Newsome et al. [51]. To specify boundary conditions for the governing equations, LARCK utilizes a generalized system of boundary conditions that can be imposed on any block face or subset of block faces. Currently, fourteen different boundary condition classes can be specified or additional boundary conditions can be input by the user.

The thermodynamics model in LARCK provides a generalized model for an arbitrary mixture of thermally perfect gases based on curve fits for specific heat and Gibbs energy. The code has an Arrhenius based finite-rate chemistry model with a generalized scheme that allows for the specification of any chosen reaction model.

A number of turbulence models have been incorporated into the LARCK code. The turbulence kinetic energy class of models include the Spallart-Allmaras model [52], the Wilcox high and low Reynolds number k-ω models (which can be used to solve the governing equations to the wall), or Wilcox’s compressible pressure gradient corrected wall matching procedure. In addition, Menter’s baseline and SST models [53,54] have been incorporated, again to solve to the wall or with Wilcox’s compressible pressure gradient corrected wall matching procedure [55]. The k-ε low Reynolds number model of Abid et al. [56] has also been included. In addition to the turbulent kinetic energy models, several algebraic Reynolds stress models have been incorporated into LARCK. These models include the algebraic Reynolds stress models of Abid et al. [57] (both the k-ε and the k-ω variants) and Givi et al. [58]. Coupling between the turbulence and chemistry fields has also been included. Gaussian or beta assumed probability density functions have been used to account for temperature variance effects on the forward and backward kinetic rate coefficients in the chemistry model [59]. Turbulence effects on the species production rates have also been accounted for by modeling the sum of the species variances using a multivariate assumed probability density function [60].

The LARCK code has been used to model individual scramjet component flow as well as the entire flow field from inlet to nozzle in a scramjet engine. It has been validated against a number of 2-D and 3-D unit problems such as the flat plate flow [61], high Mach number compression ramp flow, and Mach 3 corner flow.

3.5.4 FELISA HYP
The FELISA_HYP [18] code employs an unstructured grid algorithm specifically constructed for robust, hypersonic flow simulation. It is a finite-volume based formulation that employs an efficient edge data structure. Second-order accuracy is maintained in smooth regions using linear reconstruction following MUSCL concepts [62,63]. The Local Extremum Diminishing criteria [64] is used as a limiter near flow discontinuities. A simple, forward Euler explicit time stepping is used to relax the equations. The code is currently limited to inviscid flows. Options for equilibrium air chemistry are available.
In spite of its current limitation to inviscid flows, FELISA_HYP has proven particularly valuable in the CFD design environment because of the relatively quick grid generation capability. In the X-33 Design Phase I, the FELISA grid tools could be applied by an experienced user to generate unstructured surface and volume grids in days, as compared to a multiple-block structured grid for LAURA, which took several weeks. The actual FELISA_HYP solver is somewhat slower than the inviscid version of LAURA; however, the FELISA_HYP solver would finish several inviscid solutions on a new configuration before LAURA could even get started with a usable grid.

The FELISA_HYP inviscid flow solutions can be combined with engineering codes to extract heating data.

Even though only four codes have been discussed here, there are a number of other codes that are being used at Langley for high Mach number flows such as CFL3D [65], TLNS3D [24], PAB3D [66], OVERFLOW [67], etc. However, most of these applications have been to subsonic, transonic, and supersonic flows with only limited use to moderately high Mach number flows. This is due to the fact that these codes have no, or only limited, capability of modeling chemical processes and real gas effects present at high Mach numbers.

3.6 Applications
RANS-based computational codes are being used these days at all levels in the design and development process of advanced aerospace vehicles, but with the current state of physical and chemical modeling, their applications are reliable and quantitative only for certain flow situations. The codes can fairly well predict attached laminar and turbulent flows, including skin friction and heat transfer, over forebodies, planetary entry probes, etc. However, their predictions are, at best, qualitative for highly interacting flows at high Reynolds numbers (the type of flows encountered in high-speed inlets, combustors, corner regions, etc.). These codes play an important role in the preliminary design studies to screen out poor designs and to conduct sensitivity studies of various design parameters and their impact on incremental performance.

A number of application strategies can be used to enhance the efficiency of computational codes in an overall analysis and design process. Mesh sequencing and solution sequencing are two procedural approaches to speed up the convergence. Mesh sequencing refers to obtaining a solution on a sequence of finer grids, where each successive solution is initialized using the previous coarse grid solution. Solution sequencing refers to initializing a simulation at one trajectory point using a converged solution from a neighboring trajectory point. Solution sequencing can be used in conjunction with mesh sequencing to generate a matrix of solutions across a trajectory for a single configuration.

Another procedural strategy consists of block space marching. This strategy is useful when embedded subsonic or separated flow regions may arise in domains which are otherwise amenable to parabolized Navier-Stokes (PNS) methodology. A user probably can devise many other strategies for efficient use of computational codes to various applications.

Several application examples are now presented using some of the codes discussed earlier in the paper.

3.6.1 COMET
Aerodynamics and surface heating for the Commercial Experiment Transporter (COMET) at several points along its trajectory on return from low earth orbit were calculated with LAURA and a DSMC method [68]. The COMET module (later renamed METEOR), shown in Fig. 1, had no active control system, and relied entirely on aerodynamic forces for stability and proper orientation during its maximum heating pulse. The aerodynamic data base was used within a six degree-of-freedom trajectory code to define a splashdown footprint. The DSMC method was used to define the flow field in a transitional, rarefied regime (above 90 km); LAURA was used to define the flow field in the transitional to continuum regime (below 90 km). Wake flow had to be included prior to the peak heating point because of the large initial angle of attack. Continuum and rarefied aerodynamic predictions for lift, drag, and moment were in good agreement at 90 km. Thermochemical nonequilibrium models including 7 species for air were used down to Mach 15. Both viscous and inviscid solutions were used below Mach 15. Wake flow was included at Mach 1.5 to account for important base flow effects on aerodynamics. A matrix of 46 solutions was completed between February 14 and March 23, 1995. This matrix included 10 reacting, viscous flow solutions with wake; 13 reacting, viscous flow solutions without wake, 6 perfect gas, viscous flow solutions with wake, and 17 perfect gas, inviscid flow solutions without wake. Angles of attack varied from 0 to 90 degrees. A solution adaptive grid was employed to swing the extended grid in the wake around the body behind the base at 0 degrees to off the side at 90 degrees. Maximum job size was 71.2 MW on the C-90 and required 8.8 hours on a 72x36x64 cell domain. Actual time on the computer for this case was only 0.98 hours because of extensive use of asynchronous macrotasking relaxation. The large average concurrent CPU usage enabled fast turnaround for this large matrix of cases.

Computed results were obtained prior to initiation of the wind tunnel test program and were in excellent agreement with wind tunnel data at Mach 0, as can be seen in Fig. 2. Flight data is not available because the mission was aborted on ascent.

3.6.2 X-33 and Reusable Launch Vehicle (RLV)
Numerical simulations of hypersonic flow over preliminary configurations for a RLV and X-33, a technology demonstrator for the RLV, in support of Phase I Lockheed Martin Design are described. The simulations were executed using both chemical equilibrium and nonequilibrium gas models. Simulations were generated over six representative trajectory points for descent of the B1001 RLV configuration in order to establish traceability of aerothermodynamic design issues. Simulations were generated over five representative trajectory points for descent of the B1001A X-33 configuration. Trajectory points for simulation were chosen near peak heating and peak dynamic pressure; other points were selected on the basis of convenient anchors for Mach number and angle of attack variation. Representative surface heating, temperature and pressure distributions were provided to the design team, some examples of which are presented here. Procedures for incorporating CFD solutions into engineering code (like MINIVER [69]) format for subsequent use by the thermal design team are also discussed.

Figures 3 and 4 show the B1001 RLV configuration (reference length 1419.25 inches) and the B1001A X-33 configuration (reference length 752.2 inches), respectively, which were used
in simulations. Both configurations assume a moment center at 66 percent of the reference length behind the nose. The vehicle geometries are identical to scale from the nose to upstream of the wing (hypervator) root. B1001A is tapered more toward the base to reduce base drag as compared to its predecessor. It has body flaps on wind and lee sides that terminate at the cowl trailing edge but extend across most of the base lateral dimension. B1001 has no control surface preceding the central base region surrounding the aerospine engines. Instead, there is an expansion surface in the central region with outboard body flaps that extend past the cowl trailing edge. Parts of the body flap that extend past the trailing edge of the B1001 were not modeled in Phase I studies because body shape had already evolved to the B1001A based on wind tunnel test results.

Surface grids for B1001 RLV configuration were constructed in four sections as shown in Fig. 5. Solutions were generated in each section sequentially in a block marching mode. The first section extended from the nose to the first terminal plane approximately 10 inches upstream of the wing (hypervator) root. The grid density in the first section was 52x64 cells. The second section was constructed with 9 blocks in the circumferential direction and a total grid density of 18x116 cells. The third section was constructed with 12 blocks in the circumferential direction and a total grid density of 18x186 cells, and the final section was constructed with 23 blocks in the circumferential direction and a total grid density of 6x293 cells. In the case of B1001A X-33 configuration, surface grids were constructed in two sections as shown in Fig. 6. Solutions were generated again in each section sequentially in a block marching mode. However, in some cases, solutions were then regenerated in a fully coupled mode. The grid density in the first section was 64x64 cells, and in the grid in the second section was constructed with 12 blocks in the circumferential direction with a total grid density of 100x247 cells.

The computational aerothermodynamic analyses were focused on defining global temperature distributions around the RLV and X-33. Thermal analysis of the tanks required time dependent data in a readily accessible format as commonly provided by the MINIVER code. The required temporal resolution on the flight trajectory was much finer than the matrix of points selected for analysis by the LAURA code alone. The MINIVER code is capable of making reasonably accurate estimates of centerline heating distributions on vehicles like RLV and X-33. However, three-dimensional flow effects occurring off-centerline generally are not well approximated by it without some externally derived corrections. The necessary corrections were provided by LAURA at off-centerline locations at the times defined in the CFD matrix.

Implementation of this procedure in Phase I for both RLV and X-33 analyses occurred as follows. Heating and temperature distributions over the vehicle were generated by LAURA and compared with the windward centerline results from MINIVER. These comparisons established MINIVER as a reasonably accurate tool for the geometries and trajectories considered in the study. Off-centerline values were keyed to centerline values of laminar heating rate in a relatively dense matrix of computational planes through plots of $q_{Lam}/LamCL$ as a function of spanwise location in the plane. This data was input to MINIVER in a tabular form.

Transition to turbulence was assumed to occur for values of $Re/M_{E}$ between 250 and 300. Turbulent heating levels were computed downstream of this plane and values of $q_{turb}/q_{lam}$ were defined using earlier laminar solutions. These turbulent to laminar factors were also input into MINIVER in a tabular form. Heating at any point on the body was then predicted by MINIVER by computing the windside centerline value at the same axial location, multiplying by an appropriately interpolated value for $q_{turb}/q_{lam}$ for the spatial location on the body and temporal location along the trajectory, and applying an additional correction factor for turbulent flow if the transition criteria is exceeded.

Prediction of transition by LAURA and MINIVER along the windward centerline was in significant disagreement (LAURA predicted the threshold transition criteria to occur earlier in the trajectory than MINIVER). However, because these transition criteria have historically been derived from engineering code analysis like MINIVER, Phase I studies proceeded using MINIVER criteria. Establishment of a proper criteria is a subject of ongoing research.

Temperature maps of the vehicle, as predicted by LAURA for the 1200 s trajectory point of the RLV entry with undeflected flap, are presented in Fig. 7. These solutions were generated sequentially across four sections of the vehicle. It is seen that the highest temperatures occur near the wing root but are only slightly higher than the stagnation point temperature on the nose. A comparison of LAURA predictions of temperature and heating rates with the engineering code MINIVER along the windside centerline are shown in Fig. 8. In general, there is a reasonably good agreement.

Reference [70] provides additional simulations for these configurations with a flap deflection of 50 deg. and discusses the sensitivity of results to physical (turbulent viscosity) and numerical (grid-related) dissipation.

3.6.3 Sidewall Compression Scramjet Inlet

The LARCK code was used to simulate the flow at Mach 4 through a forward swept, sidewall compression scramjet inlet shown in Fig. 9 [71,72]. Sidewall compression, open bottom inlets with forward swept leading edges have greater mass capture and reduced cowl shock strengths as compared to inlet sidewalls with no sweep or aft sweep due to reduced mass transport. The results of the LARCK simulation are shown in Figures 10–12. Figure 10 shows static pressure contours along the vehicle body surface, the cowl surface, and a central plane located midway between these surfaces. Pressure contours along the streamwise symmetry plane are also shown. All contour values are keyed to the static to freestream pressure scale in the figure. Shock and expansion structure is clearly visible along each solution plane. Along the streamwise plane, the coalescence of sideward shocks can be seen downstream of the inlet inflow plane. The reflections of this shock structure with downstream movement can also be seen along the central horizontal plane. Identical shock structure can also be seen in the Mach contours given in Figure 11. A fairly uniform inlet outflow Mach number of 2.7 is predicted. A comparison between the predicted and measured bodyside centerline wall pressure data is given in Figure 12. The agreement between the computation and the data is excellent along the entire inlet length.
3.6.4 Mixing in Scramjet Combustor

The LARCK code was also used to study fuel-air mixing performance of the injection scheme in a scramjet combustor. The design consists of 6 interdigitated struts (3 along the top wall and 3 along the bottom wall). The mean flow Mach number at the entrance of the combustor is about 4.5. Helium is injected at the base of each strut through 3 injection ports. The Roe flux difference scheme with the van Leer limiter was used to evaluate the inviscid fluxes. The turbulence model employed was the Menter SST model [53] with a compressible, pressure gradient corrected wall matching procedure developed by Wilcox [55]. A total of approximately 8 million grid points was used to discretize the computational domain. The facility nozzle was computed separately to provide the combustor inflow conditions. The inside of the injectors were included in the computational domain, thus obviating the need of ad hoc approximations for the fuel injectant profiles. Figures 13 and 14 illustrate the resulting helium distribution and Mach number field, respectively. The helium injectors were highly underexpanded which accelerated the fuel from Mach 2.71 at the injection plane to approximately Mach 8. The high pressure flow between adjacent struts forces the flow over the strut tips, creating a counter rotating vortex pair. This vortex pair is responsible for the fuel plume roll up of the injectant closest to the top of each strut. These calculations required approximately 150 Cray C90 hours to complete.

4. RAREFIED AND FREE MOLECULAR FLOW CAPABILITY

Rarefaction effects in hypersonic flows occur over a wide spectrum of conditions ranging from low density (high altitudes) situations to relatively high density flows where the characteristic dimension is small. Examples are the aerothermodynamics of space vehicles at high altitude, the heating along leading edges at lower altitudes, and very localized aerothermal loads occurring at even lower altitudes such as that resulting from a shock on cowl lip interaction. Entry vehicles encompass the complete flow spectrum in terms of rarefaction, that is, from free molecular to continuum flows. During the higher altitude portion of entry where the flow is free molecular, gas-surface interactions are the dominant process influencing vehicle aerodynamics. For the transitional flow regime, bounded by the free molecular and continuum regimes, both gas-surface and intermolecular collisions are important in establishing vehicle aerothermodynamics.

For the transitional flow regime, the molecular mean-free path in the gas is significant when compared with either a characteristic distance over which flow properties change or when compared with the size of the object creating the flow disturbance. The flow that envelopes a vehicle will be in a nonequilibrium state, that is, one in which nonequilibrium exists among the various energy modes (translational and internal), the chemistry, and radiation for the more energetic flows. Furthermore, expansion of the forebody flow into the wake of a planetary probe or aerobrake extends to lower altitudes the conditions for which rarefaction effects are important and establishes the near wake closure and the level of heating experienced on a probe's afterbody or payload. This is particularly true of aeroassisted space transfer vehicles (ASTVs) where determination of wake closure is a critical issue because the low lift-to-drag ratio aeroshell designs impose constraints on payload configuration/spacecraft design. The issue is that the payload should fit into the wake in such a manner as to avoid the shear layer impingement and thereby minimize heating.

Transitional flows present unique difficulties for numerical simulations since the model equations used to describe continuum flows (Navier-Stokes) become deficient as the flow becomes more rarefied. A condition for the validity of the continuum approach is that the Knudsen number (ratio of the mean free path to a characteristic dimension) be small compared with unity. For low density flows, the particulate or molecular structure of the gas must be recognized. The basic mathematical model of such flows is the Boltzmann equation which presents overwhelming difficulties to computational methods for realistic flows. Consequently, the mathematical models that are readily applicable to both continuum and free molecular (collisionless) flows experience serious limitations when applied to transitional flows. Fortunately, direct simulation methods have evolved over the past 35 years that readily lend themselves to the description of rarefied flows. These developments have generally been concerned with the DSMC method. The DSMC method of Bird [73] along with many variants is the most used method today for simulating rarefied flows in an engineering context. The DSMC method takes advantage of the discrete structure of the gas and provides a direct physical simulation as opposed to a numerical solution of a set of model equations. This is accomplished by developing phenomenological models of the relevant physical events. Phenomenological models have been developed and implemented in the DSMC procedure to account for translational, thermal, chemical, and radiative nonequilibrium effects.

The basic DSMC algorithms have been more or less unchanged since 1985, while major improvements in geometry and grid generation have led to codes that can be more readily applied to complex configurations, as demonstrated by the flow simulations for the Shuttle Orbiter in 1990. Since 1990, efforts have continued to extend the functionality of DSMC for complex 3-D simulations. A substantial factor in the increased application of DSMC has been the rapid increase in computer capabilities that have resulted in major improvement over the past ten years for calculations on the fastest reasonably available computers.

With vector processing of DSMC codes being restricted to about a factor of five speedup over unvectorized codes, workstations have become the preferred platform for DSMC simulations. Advantages of the workstations are that they are generally more accessible and allow interactive control. However, really large DSMC simulations benefit from using parallel computing, and recent research [74] has shown that it is possible to perform simulations with over 100 million particles on a 400 node IBM SP2 computer while achieving a parallel efficiency of 90 percent.

With regard to physical models that are important in hypersonic calculations, the major advances during 1985–90 were in the modeling of slightly ionized flows and thermal radiation motivated by the technology issues associated with orbital transfer vehicles and the work initiated on the development of an Aeroassist Flight Experiment (AFE) vehicle. During the 1990–95 time period, the quantum vibrational model was introduced, which has not only improved the
vibrational modeling, but has allowed the integration of the chemical reaction and vibration modeling.

Many validation examples are currently available that provide results comparing DSMC solutions with experimental data. Results of these comparisons have, in general, been very favorable. Two examples of validation studies are reviewed herein that include Shuttle Orbiter aerodynamics at high altitude and the blunt body/wake studies conducted under the aegis of AGARD Fluid Dynamic Panel Working Group 18. Examples of current applications are presented describing the aerodynamics for both a generic single-stage-to-orbit vehicle and planetary capsules.

4.1 Algorithm

The DSMC method is a technique for the computer modeling of simulated molecules (atoms, molecules, ions, and electrons). The velocity components, position coordinates, and other relevant state information of these molecules are stored in the computer and are modified with time as the molecules are concurrently followed through representative collisions and boundary interactions in simulated physical space. This direct simulation of the physical processes contrasts with the general philosophy of CFD which is to obtain solutions of mathematical equations that model the processes. The computational task associated with the direct physical simulation becomes feasible when the gas density is sufficiently low. It also becomes necessary under these conditions because the Navier-Stokes equations do not provide a valid model for rarefied gases, and conventional CFD methods are unable to cope with the large number of independent variables that are involved in applications of the Boltzmann equation to realistic multi-dimensional problems. The time parameter in the simulation may be identified with real time, and the flow is always calculated as an unsteady flow. A steady flow is obtained as the large time state of an unsteady flow. There is no iterative procedure for convergence to the final solution, and, most importantly, there are no numerical instabilities. There is a requirement to collect a sufficient number of samples to reduce the statistical scatter to an acceptable level.

The two basic steps in a DSMC computation are the movement of the molecules and the pairing and selection of nearby molecules for collisions. The uncoupling of the molecular motion and collisions over small time steps and the division of the flowfield into small cells are the key computational assumptions associated with the DSMC method. The time step should be much less than the mean collision time, and a typical cell dimension should be less than the local mean free path. The sampled density is used in the procedures for establishing the collision rate, and it is desirable to have the number of simulated molecules of the order of ten to twenty per cell. There is a statistical consequence of the replacement of the extremely large number of real molecules by a very small number of simulated molecules. The statistical scatter generally decreases as the square root of the sample size, and, in order to attain a sufficiently small standard deviation, the simulations employ either time averaging for steady flows or ensemble averaging for unsteady flows.

The DSMC algorithm consists of the following basic steps: 1) move all molecules through a computational grid spanning physical space according to the product of the velocity of each molecule and a small time step and compute interactions with boundaries as required; 2) determine the cell location of each molecule; 3) on a statistical basis, compute collisions between molecules occupying the same cell; and 4) sample information on the molecules residing in each cell. Even though the large number of molecules in a real gas is replaced by a much smaller number of model molecules in a simulation, thousands to millions of model molecules are still needed, which can lead to substantial computer requirements.

DSMC codes have traditionally focused on algorithms that allow the greatest flexibility in modeling the physics rather than on algorithms that offer the most efficient computations. However, architectural differences between traditional vector supercomputers and workstations lead to different requirements for algorithmic implementation. Research [74] has shown that significant improvements are possible by using data structures optimized for specific platform architecture.

Much of the current development efforts are focused on extending the functionality of DSMC for complex 3-D simulations. These capabilities must be incorporated into software capable of running on massively parallel systems as well as engineering workstations to achieve useful "engineering" tools for the more demanding current applications and to provide the opportunity to explore problem areas that are currently being approached with the continuum-based methods. While it is unlikely that DSMC will replace the continuum-based methods, the particle-based methods may give new insight into these problems from a molecular point of view.

4.2 Physical and Chemical Modeling

4.2.1 Collision Cross-Section

The collision of two molecules can be simulated using models ranging from simple hard sphere interactions to more sophisticated models that include the potential fields and the internal quantum states. The engineering approach used in most DSMC calculations is based on the Variable Hard Sphere (VHS) model proposed by Bird [73]. This model is based on observations that it is the change in collision cross-section with the relative energy of the colliding molecules, rather than the variation in the scattering law, that is most responsible for the observed effects of the molecular model on rarefied flows. The VHS model has a well-defined cross-section and follows the classical hard sphere scattering law, but the cross-section is an inverse power law function of the relative collision energy between the colliding molecules. The model parameters for the VHS model are deduced from the species viscosity data. Koura and Matsumoto [75] extended this model by introducing the variable soft sphere (VSS) model which has an additional parameter that accounts for anisotropic scattering. This parameter is fitted by comparing the diffusion coefficient of the model gas with that of the real gas. Another recent model that accounts for the attractive part of the potential is the generalized hard sphere (GHS) model of Hash and Hassan [76]. It bears the same relationship to the Leonard-Jones class of interaction potentials as the VHS or VSS models bear to the inverse power law interaction.

4.2.2 Energy Exchange

A detailed description of molecular interactions includes internal energy exchange, chemical reactions, and for sufficiently energetic flows, thermal radiation. Such phenomena
can only be fully described using quantum mechanics. Since such an approach would be much more computationally expensive, the models implemented to deal with the inelastic aspects of real molecules are generally phenomenological in nature. The phenomenological approach is to create the simplest possible mathematical model of a process that reproduces the physically significant aspects of the real process.

The most important of these models is the Larsen-Borgnakke [77] model for the calculation of the internal energy exchange during binary collisions of polyatomic species. The essential feature of this model is that a fraction of the collisions are regarded as completely inelastic, and for these, new values of the translational and internal energies are sampled from the distributions of these quantities that are appropriate to an equilibrium gas. The remainder of the molecular collisions are regarded as elastic. The fraction of inelastic collisions can be chosen to match the real gas relaxation rate. While the model is physically unrealistic at the microscopic level, it is one of the few that satisfies the principle of detailed balancing and gives satisfactory results in numerous applications.

The recent introduction of quantum vibration models by Haas et al. [78] and Bergmann and Boyd [79] has led to a generalized Larsen-Borgnakke scheme [73] where the treatment of the vibrational states is in the form of quantum states. The introduction of the quantum vibrational model has not only improved vibration modeling but has allowed the integration of chemical reaction and vibration modeling.

4.2.3 Chemical Reactions
The procedures used to implement chemical reactions since the 1970’s are essentially extensions of the elementary collision theory of chemical physics. The binary reaction rate is obtained as the product of the collision rate for collisions with energy in excess of the activation energy and the probability of reaction (or steric factor).

The chemical data for gas phase reactions are almost always quoted in terms of macroscopic rate coefficients \( K(T) \). A form of the collision theory that is consistent with the interaction model is used to convert these temperature dependent rate coefficients of continuum theory into collisional energy dependent steric factors. The reactive cross section is the product of the steric factor and the elastic cross section. If a comprehensive data base of reactive cross sections was available, it could be incorporated directly into the simulation, but this is not the case. In addition, the use of the rate coefficients guarantees consistency with the continuum theories that become valid at the higher densities.

Modification to the Bird model to include the effects of vibrational energy promoting a reaction have been proposed by Haas and Boyd [80]. In Bird’s [73] latest model, the dissociation is closely linked with the vibrational levels of the diatomic molecules. That is, dissociation can be regarded as vibration to the level at which the bond between the atoms break. Given the vibrational relaxation rate, the Larsen-Borgnakke theory can be used with kinetic theory to derive analytical expressions for the reaction rate equations. The approach has been extended to treat recombination reactions and also exchange reactions in air. Experiments for reacting systems other than air must be conducted to determine if this approach is generally valid.

4.2.4 Thermal Radiation
The major advances during 1985–90 were in the modeling of slightly ionized flow and thermal radiation. These advances were motivated by the technology requirements to support the orbit transfer vehicle studies and the development of an Aeroscience Flight Experiment (AFE) vehicle. These flowfields are characterized by velocities of the order of 10 km/s with atmospheric encounter at high altitudes. The partial ionization of such flows is accompanied by electronic excitation and thermal radiation. Radiation from bound-bound transitions between electronic states can be significant in 10 km/s flows. The procedures used for calculating the population of electronic states are analogous to the Larsen-Borgnakke model that has proved successful for the rotational and vibrational degrees of freedom. Unlike the procedures for the rotational and vibrational modes in which each molecule is assigned a single energy or state, each molecule is assigned a distribution over all the available electronic states. This overcomes the computational problems associated with radiation from sparsely populated states. Details of the modeling developed for air species which account for molecular band and atomic transitions is described in [81].

4.3 Major Codes
The DSMC codes utilized at NASA Langley are those that have either been developed by Bird, or have a close heritage with the algorithm and modeling implemented in Bird’s codes. The primary code used for 2-D/asymmetric simulations is the G2 code of Bird [82]. The G2 program provides a flexible system for the specification of the flow geometry using a body fitted grid. This code can be applied to a wide variety of flows ranging from the flow past aerodynamic bodies and rocket plume flows to internal flows in high vacuum equipment. The time-averaged flow properties may be sampled if the flow is such that it becomes steady at large times. Alternatively, an ensemble average may be made over multiple runs of an unsteady flow.

For general three-dimensional simulation, three different codes are utilized: G3, F3, and DAC. A key feature distinguishing these codes is the surface and volume grid treatment. Of the 3-D codes, G3 is the one most closely aligned with the organization and features of G2. G3 utilizes a body-fitted grid where the surface and volume grid may be either structured or unstructured. The latter capability has been recently expanded and demonstrated by Wilmoth et al. in [83].

The F3 code developed by Bird [84] and further evolved by Rault [85] utilizes a two-level Cartesian grid to define the surface and unstructured computational cells. An advantage of the Cartesian based cells is the reduction in the time associated with molecule movement and assignment of molecules to grid cells. The unstructured cells are achieved through clustering of Cartesian elements. For low-Knudsen-number flows, a modification to the basic grid scheme has been implemented [85] where a local body-fitted mesh is used to resolve gradients normal to the wall. A recent addition to the Cartesian grid DSMC codes is an algorithm named DAC (DSMC Analysis Code) that has been developed by LeBeau [83]. DAC uses a variable-resolution Cartesian grid currently consisting of two levels of cells. The resolution of the first level of cells is
constant and is typically set based on the minimum desired flowfield sampling resolution for a given problem. To further refine the flowfield grid in areas of increased density or high gradients, each level-1 cell can have an additional level of embedded Cartesian refinement. This second level of refinement is independent for each level-1 cell. The ability to refine the flowfield grid locally allows DAC to meet the spatial resolution requirement without excessive global refinement. As with the unstructured G3 code, the surface geometry for DAC is specified as a collection of planar triangular elements which form an unstructured triangular grid. The surface grid is defined independently of the volume grid. The reduced volumes of Cartesian cells that are clipped by the surface are computed and, to minimize the computational effort required to determine molecule-surface interactions, the surface triangles are mapped to the Cartesian cells.

4.4 Validation

When careful attention is given to satisfying numerical requirements and implementing physical models appropriate to the problem, the DSMC method yields results that agree well with experiments. Examples of computations performed at the NASA Langley Research Center where comparisons have been made with experimental measurements include: surface aerothermal loads produced by shock/shock interactions resulting from an oblique shock interacting with the bow shock of a cylindrical model [86,87], shock-boundary-layer interactions induced by ramps [88] and flares [89] as they influence the extent of separation and surface quantities, the effects of rarefaction on blunt body flows and their associated wake flows [90], the aerodynamics and heating of a delta wing as a function of rarefaction, [91] high altitude Space Shuttle aerodynamics [85,92], and the aerothermodynamics [93] of a spherically blunted 50° half angle capsule flown as the Japanese orbital re-entry flight experiment (OREX). Additional details concerning the blunt body wake studies and Orbiter aerodynamics follows.

4.4.1 Blunt Body/Wake Flows

The AGARD Fluid Dynamics Panel WG 18 began in late 1991 to focus on several problem areas associated with hypersonic flows. One of the problems selected for investigation was blunt body flows and their associated wake closure which is important for entry probes and aerobrakes. A number of fundamental issues exist concerning such flows: how does the wake structure change as a function of rarefaction, what role does thermochemical nonequilibrium play in the near wake structure, and to what limits are continuum models realistic as rarefaction in the wake is progressively increased. Experiments have been conducted in five hypersonic facilities using the same model configuration: a 70° spherically blunted cone with nose radius equal to one-half the model base radius, a shoulder or corner radius equal to 5 percent of the base radius, and for the sting supported models the sting radius was one-fourth the model base radius. Figures 15 through 19 present comparisons of measured and computed results using the G2 code of Bird, demonstrating the ability of the DSMC method to simulate complex flows that span a range of conditions (nonreacting to reacting flows) in the transitional to continuum regime. Figure 15 demonstrates that good agreement is achieved between the measured [94] and calculated [90] heating rate values along the sting where the model is at zero incidence in a nonreacting Mach 20 nitrogen flow at three levels of rarefaction where the overall freestream Knudsen number, \( Kn_{inf} \), is the freestream mean free path divided by the base diameter. The model base diameter is 5 cm \((Rn = 1.25 \text{ cm})\). The measurements [94] were conducted in the SR3 facility of the CNRS, Meudon, France.

Nonintrusive electron beam fluorescence measurements of the flowfield density were also made for the two more rarefied test conditions in the CNRS experiments. Figure 16 presents a comparison of the DSMC calculation with measured values for the \( Kn_{inf} = 0.0045 \) condition. The overall quantitative features of the two data sets are similar with the exception of the expansion of the flow about the outer corner of the model and the sudden up-turn of the 0.5 density contour adjacent to the sting. The calculated density contours in the near wake show a concentrated expansion from the rewarded facing portion of the outer corner. This behavior is consistent with other DSMC calculations that have been made for this test condition as summarized in [95], both at 0° and 10° incidence. The measurements show a more diffuse expansion extending down the base of the model. Part of this discrepancy may be due, in part, to a measurement resolution issue, since the gradients in density are substantial near the surface and occur in a rather small volume. As suggested in [95], the up-turn of the measured density contours along the sting are most likely due to an increase in the cross sectional area of the sting starting 80.4 mm downstream of the forebody stagnation point of the model. The change in the sting configuration was not included in the numerical simulations.

Within the near wake, a stable vortex is calculated for each of the three flow conditions where the calculated size of the wake vortex as measured from the base plane to the wake stagnation point (the point in the wake where the separated flow reattaches and the velocity is zero) increases with decreasing \( Kn_{inf} \). The calculated behavior of the size of the wake vortex as influenced by rarefaction was later confirmed by the experiments performed in the vacuum wind tunnel V2G of the DLR, Göttingen. Experiments were made at three levels of rarefaction (bounded by the CNRS tests) in Mach 16 nitrogen flow with 5 cm base diameter models that had no afterbody sting. That is, the model was supported by three tungsten wires of 0.1 mm diameter. Patterson probe measurements [96,97] were made in the near wake to extract information concerning the molecular fluxes as a function of location and view direction. Figure 17 presents calculated and measured results for \( Kn_{inf} = 0.002 \) indicating good agreement for the wake centerline number flux, \( n \), ratioed to the freestream flux, \( n_{inf} \). The agreement is good in terms of both the extent of separation and the magnitude of the molecular fluxes.

Experiments were also made at much higher enthalpy flows where rarefaction effects were potentially present utilizing the Calspan LENS and Göttingen HEG shock tunnels. The LENS tests used sting supported models with a base diameter of 15.24 cm. One test was conducted at a \( Kn_{inf} = 0.0023 \) by operating the facility at low pressure conditions. The test by Holden et al. [98] was made in Mach 15.6 nitrogen at an enthalpy of about 5 MJ/kg. For the DSMC simulation, the freestream quantities were: \( V_0 = 32458 \text{ m/s}, \rho_0 = 2.807 \times 10^{-3}\text{ m}^{-3}, \) and \( T_0 = 103.7 \text{ K} \). For this test condition, the nitrogen gas that envelopes the test model is in thermal nonequilibrium, yet there is negligible dissociation. Measurements were made for surface heating and pressure with a concentration of instrumentation along the
sting to capture the free shear layer reattachment. As shown in Fig. 18 for the heat transfer distribution, the agreement between calculation and measurements is fairly good, both in the separated region and toward the end of the recompression process, indicating that the size of the base flow region is well predicted. With only two heat transfer measurements along the forebody, it is not possible to establish the experimental trend for heat transfer distribution.

Tests were conducted at higher enthalpy levels (10 to 23 MJ/kg) at the DLR, Göttingen, using both large (dp = 15.24 cm) and small (dp = 0.5 cm) models tested in HEG. The mini cone tests were at conditions where significant dissociation and potential rarefaction effects would be present. Measurements by Legge [99] consisted of only forebody heat transfer rates at the stagnation point and at an \( \sqrt{R_e} \) location of 0.6. The tests were in air at Mach numbers of approximately 10. Two of the test cases have been simulated with the DSMC method using a 5-species reacting air gas model. For the lower enthalpy condition

\[
\rho_{\infty} = 4.085 \times 10^{-3} \text{ kg/m}^3, \quad T_{\infty} = 489.9 \text{ K},
\]

\( \rho_{\infty} = 0.003 \), the maximum mole fraction of atomic nitrogen along the forebody was of the order of 0.01 while the value for the higher enthalpy

\[
V_{\infty} = 6075 \text{ m/s}, \quad T_{\infty} = 856.4 \text{ K}, \quad \rho_{\infty} = 1.564 \times 10^{-3} \text{ kg/m}^3, \quad \text{ and } \quad K_{\infty} = 0.0099
\]

test condition was of the order of 0.2. The calculated heating rate distributions for both cases are presented in Fig. 19 where the surface is assumed to be non-catalytic at a cold wall temperature of 300 K. Also shown are the measured values obtained using the thin wall technique. Good agreement is obtained for both test conditions in terms of distributions and absolute values. The estimated error of the heat transfer measurements of the mini cone tests was ±25 percent with the scatter shown [99] to be within ±20 percent.

These examples where one DSMC code has been applied to several experimental test conditions involving compressive forebody flows followed by rapid expansion into the near wake of a blunted cone provide enhanced confidence in the DSMC method to accurately simulate the complexities of such flows.

4.4.2 Shuttle Orbiter Aerodynamics

Several computational studies have been made comparing the results of 3-D DSMC computations with the aerodynamic data extracted from sensitive accelerometer measurements on the Orbiter during re-entry. The measurements [92] were made on orbit down to an altitude of 60 km using the Orbital Acceleration Research Experiment (OARE) which contained nano-g accelerometers along with a calibration station. Bird [84] first demonstrated good agreement with measured results at altitudes greater than 120 km using his F3 code which was later extensively evolved by Rault [85] and applied to altitudes as low as 100 km. More recent calculations (Blanchard et al. [88]) have been made by LeBeau using the DAC code and Wilmuth using an evolved version of Bird's G3 code (see [83] for a description of both codes). As shown in Fig. 20, the results from these three different 3-D codes are in excellent agreement with the flight measured normal-to-axial force ratios. Common to these calculations is the use of the variable hard sphere (VHS) intermolecular collision model and the gas-surface interaction model which was assumed to be diffuse with full thermal accommodation.

4.5 Applications

For reentry and planetary missions, analysis tools will be needed in both the continuum and the rarefied hypersonic regimes for predicting aerodynamics and heating. These analysis tools will be increasingly relied upon to provide data that historically have been obtained from ground-based and flight tests in order to reduce the overall development costs. Two current areas where analysis tools are actively supporting technology and flight missions are the Reusable Launch Vehicle (RLV) Technology Program and NASA's new decade-long program of robotic exploration of Mars.

In the rarefied regime, DSMC codes will be required for predicting vehicle aerodynamics and for predicting reaction control system (RCS) plume interactions. With the increasing computational demand that these activities will place on DSMC, the codes should be increasingly flexible and easy to use. This is particularly true for complex configurations and applications where it is necessary to couple DSMC with continuum CFD analyses such as that for RCS plume interactions.

4.5.1 DSMC Solutions for Generic SSTO Concept

The study of Wilmuth et al. [83] addressed the accuracy, efficiency and ease of use of different grid and geometry methodologies when applied to the DSMC computations about a generic single-stage-to-orbit (SSTO) concept for which continuum flowfield solutions at lower altitude flight conditions have been presented. The two DSMC codes used in this study were the G3 and DAC codes discussed earlier. As implemented in this study, G3 used an unstructured tetrahedral grid and DAC used a 2-level Cartesian grid, variable resolution approach. Common to both codes was the body geometry definition with emphasis on achieving geometric flexibility to deal with complex geometries. This was accomplished by specifying the surface as a collection of planar triangular elements which form an unstructured triangular grid.

DSMC simulations were performed for an altitude of 120 km with the following freestream conditions: \( V_{\infty} = 7818 \text{ m/s}, \quad n_{\text{in}} = 5.107 \times 10^{17} \text{ m}^{-3}, \quad \text{and } \quad T_{\infty} = 360 \text{ K}. \) For the results presented in Fig. 21, the vehicle was at 32° angle of incidence with an assumed wall temperature of 1000 K. Based on the freestream conditions and body length of 56.57 m, the overall Knudsen number was about 0.08. A Knudsen number based on mean aerodynamic cord length would be closer to unity, indicative of a more rarefied flow.

Comparisons of surface contours of heat transfer with the two codes are shown in Fig. 21 for both upper and lower surfaces. The results match extremely well both qualitatively and quantitatively. A comparison of aerodynamic results obtained with Newtonian calculations is shown in Fig. 22 for the lift-to-drag (L/D) ratio. The DSMC result with collisions enabled is higher than the free molecular value, which is indicative of transitional effects. Yet the flow is significantly rarefied as the DSMC result is much lower than the Newtonian limit.

The calculated center-of-pressure is ahead of the center of gravity for this flight condition, resulting in a positive (nose-up) pitching moment at all angles of attack for both the free molecular and DSMC results. This is in contrast to the nose-down pitching moment predicted for the Space Shuttle Orbiter.
with similar control deflections and under similar flight conditions [85]. However, the magnitude of the pitching moment predicted by DSMC is very small, and only a small force would be required to trim the vehicle at the attitudes shown.

The general conclusion from the study of Wilmoth et al. [83] was that even though the Cartesian schemes appear to offer the most advantages, it seems that no single grid methodology is universally superior for all problems. In fact, it is likely that a hybrid scheme in which a Cartesian grid is used for the bulk of the problem together with a local body-fitted grid near the body may offer significant advantages for near continuum solutions about RLV type vehicles.

4.5.2 Capsule Transitional Aerodynamics

Transitional aerodynamics becomes more critical for missions where passive aerodynamic control is relied upon for orientation and stabilization. For such missions, DSMC calculations can be used to more accurately define the transitional aerodynamics and validate the bridging relations used in trajectory analyses. Examples of recent applications of detailed computational tools for defining entry transitional aerodynamics is given in [68] for a commercial experiment transporter (COMET) reentry capsule, [100,101] for Mars Microprobes, [102,103] for the Stardust reentry capsule, and [101] for the Mars Pathfinder.

A unique aspect of the two Mars Microprobe entries is that the vehicles will encounter Mars' outer atmosphere in a random state, potentially backward and tumbling after deployment from their host Mars '98 lander. The passive reorientation requirement in the upper atmosphere is a requirement that has not been addressed by previous planetary programs. Results of DSMC calculations using the DAC code are presented in Fig. 23 for the moment coefficient about the center of gravity. These results demonstrate that the spherically blunted 45° half angle cone forebody (overall diameter of 0.35 m) followed by a hemispherical backshield centered at the center of gravity is statically unstable in the rear facing configuration for high Knudsen number conditions. Trajectory calculations show nominal reorientation occurs in the rarefied regime and the potential range of angle of attack is reduced prior to encountering continuum flow conditions.

DSMC calculations have shown that both the Mars Pathfinder (70.19° spherically blunted forebody followed by a 46.63° conical afterbody with an overall diameter of 2.65m) and the Stardust comet sample return capsule (60° spherically blunted forebody followed by a 30° conical afterbody with an overall diameter of 0.812m) experience static instabilities over the more rarefied portion of their transitional flow encounter when oriented at small angles of attack (Fig. 24). The calculated static pitching moment coefficient for Pathfinder. The trim angles for these vehicles at free molecular conditions are near 70° and 180°. For Pathfinder, trajectory simulations indicate that the gyroscopic effect of the vehicle spin is adequate for preventing unacceptable increase in vehicle attitude during the transitional portion of entry. For Stardust, however, the transitional static instability identified has required additional spin stabilization to prevent unacceptable excursions in angle of attack as the vehicle traverses the rarefied portion of Earth reentry and approaches peak heating conditions.

5. FUTURE AREAS OF RESEARCH

Advanced computational tools are playing a significantly increasing role in the analysis and design of hypersonic transportation and planetary entry systems in both continuum as well as rarefied flow regimes. Their rapid use in hypersonics has been driven primarily by necessity since there are not adequate ground-based facilities to simulate flight conditions. As discussed in the paper, a number of 3-D viscous codes are now available that include modeling for complex physical and chemical phenomena present in hypersonic flows. The solution algorithms in these codes are also becoming increasingly reliable and robust. However, the improvements in both solution algorithms and phenomenological modeling have been incremental, at best, over the last decade. It is necessary to reduce the time for the entire computational process by at least two to three orders of magnitude for routine use of these computational tools in analysis, design, and optimization.

Surface modeling and field grid generation on complex configurations continues to take the most time in the computational process. Currently, it takes anywhere from a few days to several weeks to generate a reasonable viscous grid for a real configuration. Although this is an area where significant progress has been made, further developments are required to reduce the discretization time to not more than a few hours with the use of automation and by creating grid libraries on a variety of shapes, appendages, configurations, etc. that can be called up, assembled, and suitably modified. The approaches in grid generation should be flexible and should allow quick definition and alterations of configurations. Although grid strategies in the near future will continue to use both structured and unstructured approaches, it appears that the unstructured grid approach is superior for complex configurations and takes much less time than the structured grid approach. It also allows optimum use of grid cells and their adaptation to high gradient regions. Once algorithms compatible with unstructured grids are well-developed, structured grid capability will probably no longer be required.

Convergence acceleration of the computational algorithms by an order of magnitude is another aspect of the solution process that can reduce the total time. Most algorithms being used today in the codes were developed in the seventies and early eighties and since then, only marginal progress has been made in improving their convergence. Further developments in multigrid algorithms and local preconditioning have the potential to provide this order of magnitude of improvements in convergence. There is also a need to build some intelligence in these algorithms so that they can sense the convergence problems or instabilities and automatically modify or adjust the algorithm to avoid catastrophic failure. Another feature that can accelerate the overall convergence is the ability of the algorithm to conduct local iterations in regions of slow convergence.

A significant savings in total computer time can be achieved by efficient use of parallel computers, which are fast becoming the computational platforms of choice. Of course, this requires development of algorithms that are compatible with the requirements of parallelization and that the code developer be knowledgeable of the architectural features of parallel computers to take full advantage of their capabilities. The preceding type of efficiency gains are essential in general, but for nonequilibrium, chemically reacting flows, they are almost mandatory if the computational tools for such flows are to be...
used on a routine basis in the analysis and design of hypersonic systems. Currently, a well resolved analysis of a complete scramjet engine flowpath can take anywhere from one to several hundred hours on a Cray C-90 class machine, making such analysis extremely expensive for routine use.

Advances in physical and chemical modeling for hypersonic flows are necessary to increase the reliability and robustness of the codes and to reduce uncertainty in the predictions. Most of the codes currently use simple models which were not intended to be used in hypersonic flows. New models for turbulence, onset and transition from laminar to turbulent flow, turbulence-chemical kinetics interactions, energy exchange mechanisms, and radiation are required. There is also evidence that local speed of sound, drag, and heating rates may be significantly affected by externally produced, weakly ionized flow in ways that are not fully understood or predicted by the current physical models. Some of these models, when incorporated in the governing equations, introduce new complexities which may require modifications in the solution algorithms. Inherent in the development of these models is the need of very detailed and well designed experimental databases, both mean and fluctuating, at appropriate flow conditions. These databases are essential not only to develop new or improved models, but also are required to understand the deficiencies of a model.

Code validation, especially for hypersonic turbulent, interacting, and chemically reacting flows, remains a major concern. Both local and global (at subsystem or component and overall system level) validation of the codes is required. Global evaluation of uncertainties is necessary to predict overall performance numbers because the combination of uncorrelated local and component or subsystem level uncertainties may sum up to very large global uncertainties. The complexity comes from the fact that the validation database is obtained in ground facilities having different environment (flow quality, contaminants, enthalpy, dynamic pressures, etc.) than that encountered in flight. Of course, geometric scaling and its impact on combustion chemistry, in particular, also remains an issue in code validation. Unless progress is made in developing improved models and codes are adequately validated, the codes will continue to remain tools only for preliminary design and incremental performance prediction in sensitivity studies.

The paper has focused primarily on the steady flow simulations. Unsteady flows produce their own challenges. Both efficient, time-accurate algorithms and techniques for storage, analysis, and display of large volumes of data generated in unsteady flow simulations are required. In addition, the paper discusses computational codes for aerodynamic analysis and design only. However, for multidisciplinary research, CFD codes need to be integrated with other disciplinary computational tools. For example, to study the aerothermostructural behavior of a vehicle, an aerodynamic code needs to interface and interact with material properties and structural analysis codes.

Many of the issues facing the continuum CFD with respect to surface and field grid generation are also relevant to the rarefied flow codes. Much of the current efforts in rarefied flow code development are focused on enhancing the functionality of DSMC for complex 3-D simulations, incorporating added "intelligence" to reduce user intervention, and implementing them on massively parallel computers. Physical models are also being enhanced. The quantum vibrational model that has evolved recently could be used to improve the modeling of thermal radiation due to vibrational transitions. A future area of research involves integration of electronic excitation and ionization in the same way that vibrational excitation and dissociation has been integrated. This would similarly improve the modeling of radiation due to electronic transitions.

It is now several years since the DSMC approach was shown [87] to be an effective tool for studying the Edney Type IV wave interaction that was important to the aerospace plane project at that time. Computers have become significantly faster since then and it is possible that DSMC could now be applied to study even some aspects of hypersonic propulsion. For example, the combustion process and the degree of recombination in the nozzle are among some of the critical issues and, whatever the density, DSMC could now be applied to one-dimensional models of these elements of the propulsion system.

ACKNOWLEDGEMENTS
The authors would like to acknowledge of contributions of Kay Wurster and Richard Wilmeth of NASA Langley Research Center, Jeffrey White of TA Technologies, Inc.; Steve Alter of Lockheed, Hampton; and Norma Bean of Computer Sciences Corporation. Special thanks are also due to Ms. Lori Rowland of NASA Langley Research Center, who spent many hours in preparing and formatting the paper.

REFERENCES


---

**Figure 1.** Full flow field and non-wake grid in symmetry plane of COMET (every other body-normal point shown)

**Figure 2.** Comparison of experimental results on 7% scale COMET model in Mach 6 Tunnel with computed results
Figure 3. B1001 configuration (RLV)

Figure 4. B1001A configuration (X33)
Figure 5. B1001 surface grid (RLV)

Figure 6. B1001A surface grid (X33)
Figure 7. Temperature contours on RLV for fully catalytic wall, laminar flow, at 1200 s

Figure 8. Windside centerline comparisons of LAURA and MINIVER results at the 1200 s trajectory point for the RLV
Figure 9. Forward-swept, sidewall compression inlet concept

Figure 10. Pressure contours through the inlet at Mach 4

Figure 11. Mach number contours through the inlet at Mach 4

Figure 12. Pressure comparison with experimental data on the bodyside centerline
Figure 13. Concentration contours for helium in the combustor

Figure 14. Mach number contours through the scramjet combustor
Figure 15. Measured and calculated heating rates along stinger of a 70° spherically blunted cone in Mach 20 nitrogen \((d_b = 5\, \text{cm})\).

Figure 16. Measured and calculated density contours for SR3 test in Mach 20 nitrogen \((Kn_\infty = 0.0045, d_b = 5\, \text{cm})\).

Figure 17. Measured and calculated number flux along centerline of near wake for V2G test in Mach 16.5 nitrogen \((Kn_\infty = 0.0021, d_b = 5\, \text{cm})\).

Figure 18. Measured and calculated heating rate distributions for LENS test in Mach 15.6 nitrogen \((Kn_\infty = 0.0023, d_b = 15.24\, \text{cm})\).
Figure 19. Measured and calculated heating rate distributions for HEG test in air (Case 1: $M_{\infty} = 10.1$ and $Kn_{\infty} = 0.0034$, Case 2: Mach 9.5 and $Kn_{\infty} = 0.0088$) ($d_{b} = 0.5 \, \text{cm}$).

Figure 20. DSMC comparisons with Shuttle flight data.

Figure 21. Heat transfer contours (W/m$^2$) on upper and lower surfaces of generic single-stage-to-orbit model ($V_{\infty} = 7.8 \, \text{km/s at 120 km altitude}$).
Figure 22. Lift-to-drag results for generic SSTO vehicle computed with tetrahedral code.

Figure 23. Computed pitching moment about C.G. for Mars Microprobe ($d_p = 0.35$ m).

Figure 24. Computed pitching moment about C.G. for Mars Pathfinder ($d_p = 2.65$ m).