

Numerical Methods for Reacting Gas Flow Simulations

S. van Veldhuizen¹, C. Vuik¹, and C.R. Kleijn²

¹ Delft University of Technology, Delft Institute of Applied Mathematics and J.M. Burgers Center, Mekelweg 4, 2628 CD Delft, The Netherlands
`{s.vanveldhuizen, c.vuik}@tudelft.nl`

² Delft University of Technology, Department of Multi Scale Physics and J.M. Burgers Center, Prins Bernhardlaan 6, 2628 BW Delft, The Netherlands
`c.r.kleijn@tudelft.nl`

Abstract. In this study various numerical schemes for simulating 2D laminar reacting gas flows, as typically found in Chemical Vapor Deposition (CVD) reactors, are proposed and compared. These systems are generally modeled by means of many stiffly coupled elementary gas phase reactions between a large number of reactants and intermediate species. The purpose of this study is to develop robust and efficient solvers for the stiff heat-reaction system. The velocities are assumed to be given. For non-stationary CVD simulation, an optimal combination in terms of efficiency and robustness between time integration, nonlinear solvers and linear solvers has to be found. Besides stability, which is important due to the stiffness of the problem, the preservation of non-negativity of the species is crucial. It appears that this extra condition on time integration methods is much more restrictive towards the time-step than stability.

1 Introduction

In Chemical Vapor Deposition (CVD) literature, and also other reactive flow literature, one is usually looking for the steady state solution of the species equations (1). The usual procedure to find this steady state solution is to perform a (damped/relaxed) Newton iteration with an (arbitrary) initial solution. Hopefully, the Newton iteration converges to the steady state. If this is not the case one performs some (artificial) time stepping in order to find a better initial solution for the next Newton iteration. In this paper we present suitable time integration methods for stiff problems. Furthermore, we compare these integration methods by their performance, in terms of efficiency.

In our research we are not looking for the steady state solution only, but we also want the transient solution. In order to simulate this transient, we have to use a time integration method that can handle stiff problems.

2 Model for CVD Simulation

The mathematical model describing the CVD process consists of a set of PDEs with appropriate boundary and initial conditions, which describe the gas flow, the transport of energy, the transport of species and reactions in the reactor.

The gas mixture in the reactor is assumed to behave as a continuum. The gas flow in the reactor is assumed to be laminar. Since no large velocity gradients appear in CVD gas flows, viscous heating due to dissipation will be neglected. We also neglect the effects of pressure variations in the energy equation. The composition of the N component gas mixture is described in terms of the dimensionless mass fractions $\omega_i = \frac{\rho_i}{\rho}$, $i = 1, \dots, N$, having the property $\sum_{i=1}^N \omega_i = 1$. The transport of mass, momentum and heat are described respectively by the continuity equation, the Navier-Stokes equations and the transport equation for thermal energy expressed in terms of temperature T . See for instance [4, 6].

We assume that in the gas-phase K reversible reactions take place. For the k^{th} reaction the *net* molar reaction rate is denoted as R_k^g ($\frac{\text{mole}}{\text{m}^3 \cdot \text{s}}$). For an explicit description of the net molar reaction rate, we refer to [4, 6]. The mass diffusion flux is decomposed into concentration diffusion and thermal diffusion. In this study we describe ordinary diffusion in terms of effective diffusion coefficients \mathbb{D}'_i , such that we obtain

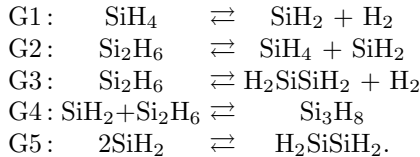
$$\frac{\partial(\rho\omega_i)}{\partial t} = -\nabla \cdot (\rho\mathbf{v}\omega_i) + \nabla \cdot (\rho\mathbb{D}'_i\nabla\omega_i) + \nabla \cdot (\mathbb{D}_i^T\nabla(\ln T)) + m_i \sum_{k=1}^K \nu_{ik} R_k^g, \quad (1)$$

where \mathbb{D}_i^T the multi-component thermal diffusion coefficient for species i .

The main focus of our research is on efficient solvers for the above species equation(s) (1). Typically the time scales of the slow and fast reaction terms differ orders of magnitude from each other, and from the time scales of the diffusion and advection terms, leading to extremely stiff systems.

Simplified System. Since our research focuses on solving the species equations (1), we will only solve the coupled system of N species equations, where N denotes the number of gas-species in the reactor. The velocity field, temperature field, pressure field and density field are computed via another simulation package developed by Kleijn [5]. Other simplifications are the omission of surface reactions and thermal diffusion.

We consider a CVD process, which is a simplification of the CVD process considered in [5], that deposits silicon Si from silane SiH_4 . The gas-mixture consists of 7 species and the reaction mechanism



The studied reactor configuration is illustrated in Figure 1. As computational domain we take, because of axisymmetry, one half of the $r-z$ plane. The pressure in the reactor is 1 atm. From the top a gas-mixture, consisting of silane, with mass fraction $f_{\text{in},\text{SiH}_4} = 0.001$, and helium (the rest), enters the reactor with a uniform temperature $T_{\text{in}} = 300$ K and a uniform velocity u_{in} . At a distance of

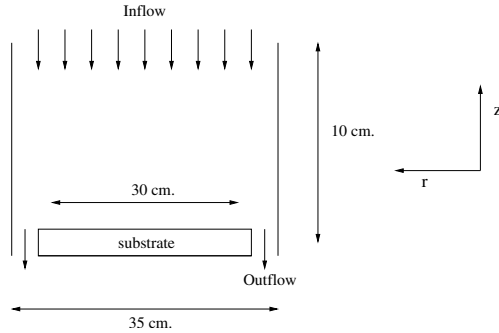


Fig. 1. Reactor geometry

10 cm. below the inlet a non rotating susceptor with temperature $T = 1000$ K and a diameter 30 cm. is placed.

We emphasize that this test-problem is not representing a practical process, but representing its computational problems. Further details on the test-problem can be found in [6].

3 Properties of Numerical Methods for Solving the Species Equations

As seen in the previous section the species eqns. (1) are PDEs of the advection-diffusion-reaction type. In order to have a unique solution appropriate boundary conditions and initial values have to be chosen.

To approximate the solution we use the Method of Lines (MOL), i.e., we first discretize in space, resulting into the ODE system

$$w'(t) = F(t, w(t)), \quad 0 < t \leq T, \quad w(0) \text{ given.} \quad (2)$$

The next step is to integrate the ODE system (2) with an appropriate time integration method. We remark that the stiff reaction terms in CVD motivates to integrate parts of $F(t, w(t))$ implicitly. In general, due to the nonlinearities in the reaction term, (huge) nonlinear systems have to be solved.

The topic of this research is to find the best combination of time integration, nonlinear and linear solvers in terms of efficiency. Note that if the computational cost of one time step is (very) expensive, then a time integration method that needs more, but computational cheaper, time steps is better in terms of efficiency.

Besides the efficiency criteria, also some other properties are desired for the numerical methods. As already mentioned in Section 2 the system of species equations is stiff. Following [2], we say that stiffness indicates a class of problems for which implicit methods perform (much) better than explicit methods. The eigenvalues of the Jacobian $\frac{\delta f}{\delta y}$ play certainly a role in this decision, but quantities such as the dimension of the system and the smoothness of the solution are also important.

Positivity. A natural property for mass fractions is that they are non-negative. As a consequence, it should also hold for the mathematical model, spatial discretization and time integration of the process. While the first one is obvious, the latter two should not introduce any (small) negative components causing blow up of the solution. It appears that this extra condition on time integration methods is much more restrictive towards the time step than stability. We remark that positivity for spatial discretization can be assured by locally first order upwinding.

An ODE system $w'(t) = F(t, w(t))$, $t \geq 0$, is called positive if $w(0) \geq 0$ implies $w(t) \geq 0$ for all $t > 0$. It is easy to prove that linear systems $w'(t) = Aw(t)$ are positive if and only if $a_{ij} \geq 0$ for $i \neq j$. See [7]. For general nonlinear semi-discretizations $w'(t) = F(t, w(t))$, it appears that unconditional positivity is a very restrictive requirement. Suppose that $F(t, w(t))$ satisfies the condition:

Condition 1. *There is an $\alpha > 0$, with α as large as possible, such that $\alpha\tau \leq 1$ and*

$$v + \tau F(t, v) \geq 0 \quad \text{for all } t \geq 0, v \geq 0. \quad (3)$$

Application of Euler Forward to the nonlinear system $w'(t) = F(t, w(t))$ gives

$$w_{n+1} = w_n + \tau F(t_n, w_n). \quad (4)$$

Provided that $w_n \geq 0$, Condition 1 guarantees positivity for w_{n+1} computed via Euler Forward (4). Furthermore, assume that $F(t, w(t))$ also satisfies:

Condition 2. *$F(t, v)$ is continuously differentiable and*

$$\|I - \tau J_F(t, v)\| \leq C, \quad \text{for any } v \in \mathbb{R}^n, t \geq 0 \text{ and } \tau > 0, \quad (5)$$

whereby C is a positive constant, I the identity matrix and $J_F(t, v)$ the Jacobian matrix of derivatives of $F(t, v)$ with respect to v .

In [2] it has been proven that Condition 1 and 2 imply positivity for Euler Backward for any step-size τ . However, in practice the solutions of the resulting nonlinear systems have to be approximated, which might introduce small negative components. In the case the negative components of the solution are the result of rounding errors, then it is justified to set them to zero. In the case one has negative components in the solution as consequence of the nonlinear (Newton) solver, then the most common method to avoid negative concentrations is *clipping*. Clipping has the disadvantage that mass is not longer preserved. In practice it is desired not to use clipping. In our experience, the implementation of variable time step size algorithm, see [6], avoids negative concentrations without the use of clipping.

We conclude this section with the claim that Euler Backward is the only time integration that is unconditionally positive. For a proof we refer to [1]. This means that for any higher order (implicit) time integration method a time step criterion is needed to ensure preservation of non-negativity.

4 Suitable Time Integration Methods (TIM)

In this section we briefly present integration methods that are suitable, from a theoretic point of view, for the time integration of the species equations. More comprehensive descriptions are given in [2, 6]. At the end of this section we will also make some remarks on the linear and nonlinear solvers.

From the previous section it is clear that the Euler Backward method is a suitable method to perform time integration. It has the advantage of being unconditionally positive. Disadvantages are the first order consistency and the probably high computational costs for one time step. The latter is due to the fact that the succeeding approximations are computed in an implicit manner.

4.1 Time Integration Methods

We will discuss a selection of time integration methods that have good properties in both stability and positivity, or TVD.

Rosenbrock Methods. The two stage Rosenbrock method

$$\begin{aligned} w_{n+1} &= w_n b_1 k_1 + b_2 k_2 \\ k_1 &= \tau F(w_n) + \gamma \tau \mathbf{A} k_1 \\ k_2 &= \tau F(w_n + \alpha_{21} k_1) + \gamma_{21} \tau \mathbf{A} k_1 + \gamma \tau \mathbf{A} k_2, \end{aligned} \quad (6)$$

with $\mathbf{A} = F'(w_n)$ is the Jacobian matrix of F , and $b_1 = 1 - b_2$, $\alpha_{21} = \frac{1}{2b_2}$ and $\gamma_{21} = -\frac{\gamma}{b_2}$, is interesting. The method is of order two for arbitrary γ as long as $b_2 \neq 0$. The stability function is given as

$$R(z) = \frac{1 + (1 - 2\gamma)z + (\gamma^2 - 2\gamma + \frac{1}{2})z^2}{(1 - \gamma z)^2}. \quad (7)$$

For $\gamma \geq \frac{1}{4}$ the method is unconditionally stable. For $\gamma_+ = 1 + \frac{1}{2}\sqrt{2}$, we have the property that $R(z) \geq 0$, for all negative real z . For diffusion-reaction problems, which have negative real eigenvalues, this property ensures positivity of the solution. It appears that the second order Rosenbrock method performs quite well with respect to the positivity property, as has been shown in [8]. In [8] it is conjectured that the property $R(z) \geq 0$ for all negative real z plays a role.

Backward Differentiation Formulas (BDF). The k -step BDF methods are implicit, of order k and defined as

$$\sum_{j=0}^k \alpha_j w_{n+j} = \tau F(t_{n+k}, w_{n+k}), \quad n = 0, 1, \dots, \quad (8)$$

which uses the k past values w_n, \dots, w_{n+k-1} to compute w_{n+k} . Remark that the most advanced level is t_{n+k} instead of t_{n+1} . The 1-step BDF method is Backward Euler. The 2-step method is

$$\frac{3}{2}w_{n+2} - 2w_{n+1} + \frac{1}{2}w_n = \tau F(t_{n+2}, w_{n+2}), \quad (9)$$

and the three step BDF is given by

$$\frac{11}{6}w_{n+3} - 3w_{n+2} + \frac{3}{2}w_{n+1} - \frac{1}{3}w_n = \tau F(t_{n+3}, w_{n+3}). \quad (10)$$

Remark 1. A disadvantage of linear multi-step methods is that the first $k - 1$ approximations cannot be computed with the linear k -step scheme. To compute the first $(k - 1)$ approximations, one could use for the first step a BDF 1-step method, for the second approximation a BDF 2-step method, ... and for the $(k - 1)^{\text{st}}$ approximation a BDF $(k - 1)$ -step scheme.

For the 2-step BDF method we obtain positivity, under Conditions 1 and 2, of $w'(t) = F(t, w(t))$ whenever $\alpha\tau \leq \frac{1}{2}$, provided that w_1 is computed from w_0 by a suitable starting procedure, i.e., w_1 has been computed such that $w_1 \geq 0$ holds. For a derivation we refer to [2, 6].

IMEX Runge-Kutta Chebyshev Methods. The second order Runge-Kutta Chebyshev method is given as

$$\begin{aligned} w_{n0} &= w_n, \\ w_{n1} &= w_n + \tilde{\mu}_1 \tau F(t_n + c_0 \tau, w_{n0}), \\ w_{nj} &= (1 - \mu_j - \nu_j)w_n + \mu_j w_{n,j-1} + \nu_j w_{n,j-2} + \tilde{\mu}_1 \tau F(t_n + c_{j-1} \tau, w_{n,j-1}) + \tilde{\nu}_j \tau F(t_n + c_0 \tau, w_{n0}), \quad j = 1, \dots, s \\ w_{n+1} &= w_{ns}. \end{aligned} \quad (11)$$

The coefficients $\omega_0, \omega_1, b_j, c_j, \tilde{\mu}_j, \dots$ can be found in [6]. In Figure 2 its stability region is given. The parameter $\beta(s)$ moves to $-\infty$ when the number of stages s increases.

The IMEX extension of the above scheme is as follows. Suppose we have an ODE system $w'(t) = F(t, w(t))$, where $F(t, w(t))$ can be split as

$$F(t, w(t)) = F_E(t, w(t)) + F_I(t, w(t)). \quad (12)$$

In (12) the term $F_I(t, w(t))$ is the part of F which is (supposed to be) too stiff to be integrated by an explicit Runge-Kutta Chebyshev method. Obviously, the term $F_E(t, w(t))$ is the moderate stiff part of F that can be integrated in an explicit manner using RKC methods. The first stage of (11) becomes in the IMEX-RKC scheme

$$w_{n1} = w_n + \tilde{\mu}_1 \tau F_E(t_n + c_0 \tau, w_{n0}) + \tilde{\mu}_1 \tau F_I(t_n + c_1 \tau, w_{n1}), \quad (13)$$

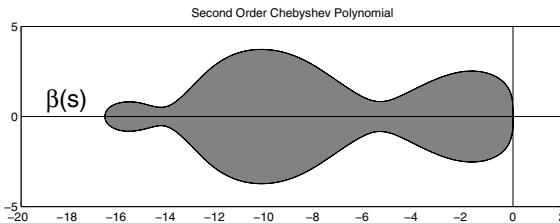


Fig. 2. Stability region of (11) with $s = 5$

with $\tilde{\mu}_1$ as defined before. Note that the highly stiff part of F is treated implicitly. The other $(s - 1)$ subsequent stages of (11) will be modified in a similar way.

With respect to stability of this IMEX extension of (11) we remark that the implicit part is unconditionally stable, whereas the stability condition for the explicit part remains unchanged.

4.2 Nonlinear and Linear Solvers

In all suitable TIM nonlinear systems $F(x) = 0$, $x \in \mathbb{R}^n$ have to be solved. The Newton iteration is, with its second order convergence, an obvious choice. The disadvantage of having local convergence will disappear if one uses a line-search algorithm, such that

$$\|F(x^{k+1})\| \leq \|F(x^k)\| \quad k = 0, 1, 2, \dots, \quad (14)$$

for some norm in \mathbb{R}^n . More background information can be found in [3, 6].

In the Newton iteration linear systems have to be solved. In most 2D applications direct solvers like LU factorization are still applicable. To reduce the amount of work one usually re-arranges the unknowns, in order to reduce the bandwidth of the matrix. Also in our case it is possible to reduce the bandwidth of the Jacobian considerably. The way to do this is described in [7].

For 3D problems direct solvers (LU factorization) are no longer applicable. To approximate the solution of the linear systems one has to switch to iterative linear solvers like, for instance, Krylov Subspace methods.

5 Numerical Results

In this section we compare the performance of the TIM of Section 4.1 for solving the species equations. If necessary, nonlinear systems will be solved by the full Newton iteration. Linear systems will be solved using the LU factorization with rearranging, as mentioned in Section 4.2. Since the solutions differ not much, we are only interested in the performance of the TIM.

For a comparison between the workloads of the various TIM, we look to the amount of CPU time, the number of time steps and the workload per time step it takes to reach steady state. We say that steady state is reached when the norm of the right hand side of (1) is below a threshold, in our case 10^{-4} , and the fluxes at the boundaries are constant. The threshold is equal to the accuracy

Table 1. Workloads of various TIM (performed on a Pentium IV 3.2 GHz)

TIM	CPU time	# time steps	Average CPU time / time step
Euler Backward	600 CPU sec	163	3 CPU sec
2 nd order Rosenbrock	375 CPU sec	100	3 CPU sec
BDF-2	600 CPU sec	63	8 CPU sec
IMEX-RKC(5)	> 5000 CPU sec	> 500	25 CPU sec

of the solution of the system of Navier Stokes eqns., pressure eqn., etc. For the different TIM mentioned in Section 4.1 the workloads are given in Table 1.

6 Conclusions

Based on Table 1, we conclude that for this 2D test-problem Rosenbrock is the cheapest TIM to solve (1). The ‘bad’ performance of the IMEX RKC scheme is due to that per time step more than one nonlinear system has to be solved. Although these nonlinear systems are cheaper to solve, it did not pay off in this 2D test case. However, this property can become interesting in 3D simulations.

References

1. C. BOLLEY AND M. CROUZEIX, *Conservation de la Positivité Lors de la Discrétisation des Problèmes d’Évolution Paraboliques*, RAIRO Anal. Numer. 12, pp. 237-245, (1973)
2. W. HUNSDORFER AND J.G. VERWER, *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*, Springer Series in Computational Mathematics, 33, Springer, Berlin, (2003)
3. C.T. KELLEY, *Solving Nonlinear Equations with Newton’s Method*, Fundamentals of Algorithms, SIAM, Philadelphia, (2003)
4. C.R. KLEIJN, *Transport Phenomena in Chemical Vapor Deposition Reactors*, PhD thesis, Delft University of Technology, (1991)
5. C.R. KLEIJN, *Computational Modeling of Transport Phenomena and Detailed Chemistry in Chemical Vapor Deposition- A Benchmark Solution*, Thin Solid Films, 365, pp. 294-306, (2000)
6. S. VAN VELDHUIZEN, *Efficient Solution Methods for Stiff Systems of Advection-Diffusion-Reaction Equations*, Literature Study, Technical Report at the Delft University of Technology, TWA-05-05 , (2005)
7. S. VAN VELDHUIZEN, *Time Integration for Reacting Gas Flow Simulations*, Technical Report at the Delft University of Technology, to appear, (2006)
8. J.G. VERWER, E.J. SPEE, J.G. BLOM AND W. HUNSDORFER, *A Second-Order Rosenbrock Method Applied to Photochemical Dispersion Problems*, SIAM Journal on Sci. Comp., 20, pp.1456-1480, (1999)