

ASC Report No. 27/2011

IMEX Methods for the ANTARES Code

N. Happenhofer, O. Koch, F. Kupka

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-04-9

Most recent ASC Reports

- 26/2011 *Michael Dreher, Angar Jüngel*
Compact families of piecewise constant functions in $L_p(0, T; B)$
- 25/2011 *Jens Geier, Anton Arnold*
WKB-based schemes for two-band Schrödinger equations in the highly oscillatory regime
- 24/2011 *Markus Aurada, Michael Ebner, Michael Feischl, Samuel Ferraz-Leite, Petra Goldenits, Michael Karkulik, Markus Mayr, Dirk Praetorius*
HILBERT-A MATLAB Implementation of Adaptive 2D-BEM
- 23/2011 *JinMyong Kim, Anton Arnold, Xiaohua Yao*
Estimates for a class of oscillatory integrals and decay rates for wave-type equations
- 22/2011 *Markus Aurada, Michael Feischl, Michael Karkulik, Dirk Praetorius*
Adaptive coupling of FEM and BEM: Simple error estimators and convergence
- 21/2011 *Michael Feischl, Michael Karkulik, Jens Markus Melenk, Dirk Praetorius*
Residual a-posteriori error estimates in BEM: Convergence of h-adaptive algorithms
- 20/2011 *Markus Aurada, Michael Feischl, Michael Karkulik, Dirk Praetorius*
Adaptive coupling of FEM and BEM: Simple error estimators and convergence
- 19/2011 *Petra Goldenits, Dirk Praetorius, Dieter Suess*
Convergent geometric integrator for the Landau-Lifshitz-Gilbert equation in micromagnetics
- 18/2011 *M. Aurada, M. Feischl, M. Karkulik, D. Praetorius*
A Posteriori Error Estimates for the Johnson-Nédélec FEM-BEM Coupling
- 17/2011 *Michael Feischl, Marcus Page, Dirk Praetorius*
Convergence of adaptive FEM for elliptic obstacle problems

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: <http://www.asc.tuwien.ac.at>
FAX: +43-1-58801-10196

ISBN 978-3-902627-04-9

© Alle Rechte vorbehalten. Nachdruck nur mit Genehmigung des Autors.



IMEX Methods for the ANTARES Code*

N. Happenhofer, O. Koch, and F. Kupka[†]

June 26, 2011

Contents

1	Strong Stability of IMEX Runge–Kutta Methods	1
2	Strongly Stable IMEX Schemes from the Literature	4
2.1	An SSP2(2,2,2) Method	4
2.2	An SSP2(3,3,2) Method	8
2.2.1	The Explicit SSP(3,2) Scheme	10
2.3	An SSP3(3,3,3) Method	14
3	The ANTARES code for astrophysical simulations	17
3.1	Simulation of Semiconvection	19
3.2	Implementation in ANTARES	19
3.3	Simulations	22
4	Comparisons	22

1 Strong Stability of IMEX Runge–Kutta Methods

We consider the ODE initial value problem

$$\dot{y}(t) = F(y(t)) + G(y(t)), \quad y(0) = y_0, \quad (1)$$

where we assume that the vector fields F and G , resulting from the spatial semi-discretization of a partial differential equation, have different stiffness, such that it seems promising to use different Runge–Kutta formulae for the treatment of the two vector fields. We can imagine this situation to arise if F is the discretization of a (linear) differential operator, and G a smooth nonlinear vector field. For this type of problems, *partitioned Runge–Kutta schemes* [9] are popular. These are also called *additive Runge–Kutta schemes*.

An s -stage partitioned Runge–Kutta method characterized by coefficient matrices $A = (a_{i,j})$ and $\tilde{A} = (\tilde{a}_{i,j})$ defines one step $y_{\text{old}} \rightarrow y_{\text{new}}$ by

$$y_i = y_{\text{old}} + \Delta t \sum_{j=1}^s a_{i,j} F(y_j) + \Delta t \sum_{j=1}^s \tilde{a}_{i,j} G(y_j), \quad i = 1, \dots, s, \quad (2)$$

$$y_{\text{new}} = y_{\text{old}} + \Delta t \sum_{j=1}^s b_j F(y_j) + \Delta t \sum_{j=1}^s \tilde{b}_j G(y_j). \quad (3)$$

If $a_{i,j} = 0$ for $j \geq i$, the method is referred to as an *implicit–explicit (IMEX)* method.

These methods have first been investigated with respect to the TVD property (which more generally is referred to as *strong stability preserving (SSP)* or *monotonicity* when norms other than the total variation norm or even sublinear functionals are considered), in the context of hyperbolic systems with relaxation, where $G = \frac{1}{\varepsilon} \hat{G}$, $\varepsilon \ll 1$, in [25]. Ibidem, the common specification for strongly stable IMEX methods is introduced. An IMEX method is referred to as ‘SSP $k(s, \sigma, p)$ ’ when it has the following properties: k is the order of the method in the stiff limit $\varepsilon \rightarrow 0$, which is characterized by the coefficients for the explicit part. The latter must necessarily be SSP and is referred to as the *asymptotically SSP scheme*. s is the number of stages in the implicit scheme and σ the number of stages in the explicit scheme. p is the global order of the resulting combined method.

*The authors would like to acknowledge financial support of this work by the Austrian Science Fund FWF, projects P20793 and P21742-N16.

[†]All: Faculty of Mathematics, University of Vienna, Nordbergstraße 15, A–1090 Wien, Austria

It is essential to observe that if the implicit scheme characterized by $\tilde{A} = (a_{i,j})$ is a *diagonally implicit Runge–Kutta (DIRK)* method, then the explicit part is evaluated only once in each stage, providing the desired computational advantage [25]. The analysis given in [25] is valid only for $\varepsilon \ll 1$ [14]. However, several useful examples of strongly stable IMEX Runge–Kutta methods are given, see Section 2 below.

[14] develops a comprehensive theory of strong stability preserving additive Runge–Kutta schemes which extends the concepts for standard Runge–Kutta methods in a natural way:

Let $\frac{1}{\rho}, \frac{1}{\tilde{\rho}}$ be the step-size coefficients for monotonicity of the explicit Euler method for the vector fields F and G , respectively. We define the *region of absolute monotonicity*

$$\mathcal{R}(A, \tilde{A}) = \{(r, \tilde{r}) \in \mathbb{R}^2 : (A, \tilde{A}) \text{ is absolutely monotonous on } [-r, 0] \times [-\tilde{r}, 0]\}, \quad (4)$$

where the absolute monotonicity at a point (r_0, \tilde{r}_0) is characterized by algebraic relations for the matrices A, \tilde{A} . The boundary in the first quadrant, $\partial\mathcal{R}(A, \tilde{A}) \cap \{(r, \tilde{r}) : r, \tilde{r} \geq 0\}$, is denoted as the *curve of absolute monotonicity*. The significance of the region $\mathcal{R}(A, \tilde{A})$ is expressed in the following theorem [14]:

Theorem 1.1 *Let (A, \tilde{A}) be absolutely monotonous at $(-r, -\tilde{r})$ with step-size coefficients $\rho, \tilde{\rho}$. Then for $h \leq \min\left\{\frac{r}{\rho}, \frac{\tilde{r}}{\tilde{\rho}}\right\}$, diminishing of the norm holds,*

$$\|y_i\| \leq \|y_{\text{old}}\|, \quad \|y_{\text{new}}\| \leq \|y_{\text{old}}\|.$$

[15] gives order barriers for strongly stable additive Runge–Kutta methods similarly to [20]. The order of an additive Runge–Kutta method (A, \tilde{A}) is bounded by the orders of A and \tilde{A} , respectively. This implies for IMEX methods the order barrier $p \leq 4$ [20]. Moreover, [15] gives a simple algebraic criterion for a nontrivial region of absolute monotonicity in terms of incidence matrices of A, \tilde{A} . Finally, some examples of strongly stable IMEX Runge–Kutta methods are given, see Section 2 below.

The stability region of IMEX Runge–Kutta methods is defined in [2, 1] via the test equation where in (1)

$$F(u) = i\beta u, \quad G(u) = \alpha u, \quad \alpha \leq 0 < \beta. \quad (5)$$

For a given time step Δt , let

$$x = \alpha\Delta t, \quad y = \beta\Delta t, \quad z = x + iy,$$

and recast the scheme (2)–(3) applied to (5) as

$$y_{\text{new}} = y_{\text{old}} + z \sum_{j=1}^s b_j y_j =: R(z)y_{\text{old}}, \quad (6)$$

$$y_i = \frac{1}{1 - x\tilde{a}_{ii}} \left((1 + iy a_{i+1,1}) y_{\text{old}} + \sum_{j=1}^{i-1} (x\tilde{a}_{i,j} + iy a_{i+1,j+1}) y_j \right), \quad (7)$$

where throughout our discussion of IMEX schemes, we assume for simplicity $b_j = \tilde{b}_j$ (which is indeed satisfied for our methods in Section 2). The additional assumptions underlying formulas (6)–(7) are that the number of stages in the explicit scheme is one larger than the number of stages in the implicit scheme, and that the method is applied to the test equation with

$$f(y) = i\beta y \quad \text{and} \quad g(y) = \alpha y. \quad (8)$$

We will perform the corresponding analysis of the stability function

$$y_{\text{new}} = R(z)y_{\text{old}} \quad (9)$$

for the methods we consider in Section 2.

Since the number of stages is equal for the explicit and the implicit schemes we use, the formalism above is awkward to apply to the methods in our focus. Thus, to investigate the stability region of interest in treating the test equation (8), we will compute $R(z)$ directly from the definition (9).

Furthermore, to check the correctness of our realization of the IMEX methods, we will compute the empirical convergence orders of the methods. We solve the non-linear test problem

$$y'(t) = (1 + \sin(y(t))) + (y^2(t) - \sin(y(t))), \quad y(0) = 0, \quad (10)$$

with the known exact solution $y(t) = \tan(t)$. Tables 1, 4, and 10 below give the stepsizes Δt , exact errors and empirical convergence order at $t = 1.3$ computed as

$$\text{order} = -\log \left(\frac{|y_N^h - y(1.3)|}{|y_{2N}^{h/2} - y(1.3)|} \right) / \log(2),$$

for the three IMEX schemes, respectively, when applied to (10). The last column gives the error constant C . Note that as required this converges to a constant independent of Δt . The error constant is smallest for the scheme (13) and largest for the third order method (23).

Finally, we wish to investigate the dissipativity of the time integrators in conjunction with suitable space discretizations. The stability function defined above does not lend itself readily to a dissipativity analysis of the ilk of [31, 18]. Accordingly, [31] gives a justification for considering only the diffusion term in (5) since the advection term becomes negligible asymptotically. We will thus investigate the dissipativity of the implicit scheme specified by \tilde{A} . To this end, we apply the spatial discretizations L_h in our focus to the heat equation

$$u_t = bu_{xx},$$

and associate for the spatial discretization $u_{j\pm k} \leftrightarrow e^{\pm i\theta k}$. Finally, we compute the *amplification factor* $g(\theta, \mu) = R((L_{\Delta x}u)_j)$, with $\mu = b\frac{\Delta t}{(\Delta x)^2}$. The spatial discretizations which were found to show a dissipative behaviour in [18] are the three-point scheme (the upper index refers to the time step)

$$(u_i^n)_{xx} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2} + O((\Delta x)^3), \quad (11)$$

and the fourth order stencil

$$(u_i^n)_{xx} = \frac{-u_{i+2}^n + 16u_{i+1}^n - 30u_i^n + 16u_{i-1}^n - u_{i-2}^n}{12(\Delta x)^2}. \quad (12)$$

2 Strongly Stable IMEX Schemes from the Literature

The strongly stable IMEX schemes listed in the following all have DIRK (Diagonally Implicit Runge-Kutta) methods as the implicit scheme, which ensures that the stages can be solved successively, and the explicit part only has to be evaluated once.

2.1 An SSP2(2,2,2) Method

[25] gives an IMEX SSP2(2,2,2) method with nontrivial region of absolute monotonicity ($\gamma = 1 - \frac{1}{\sqrt{2}}$):

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline A & \frac{1}{2} & \frac{1}{2} \end{array} \quad \begin{array}{c|cc} \gamma & \gamma & 0 \\ 1-\gamma & 1-2\gamma & \gamma \\ \hline \tilde{A} & \frac{1}{2} & \frac{1}{2} \end{array} \quad (13)$$

The method is L -stable with $\mathcal{R}(A) = 1$, $\mathcal{R}(\tilde{A}) = 1 + \sqrt{2}$, and

$$\mathcal{R}(A, \tilde{A}) = \{(r, \tilde{r}) : 0 \leq r \leq 1, 0 \leq \tilde{r} \leq \sqrt{2}(1-r)\},$$

see [14]. A plot of $\mathcal{R}(A, \tilde{A})$ is given in Figure 1.

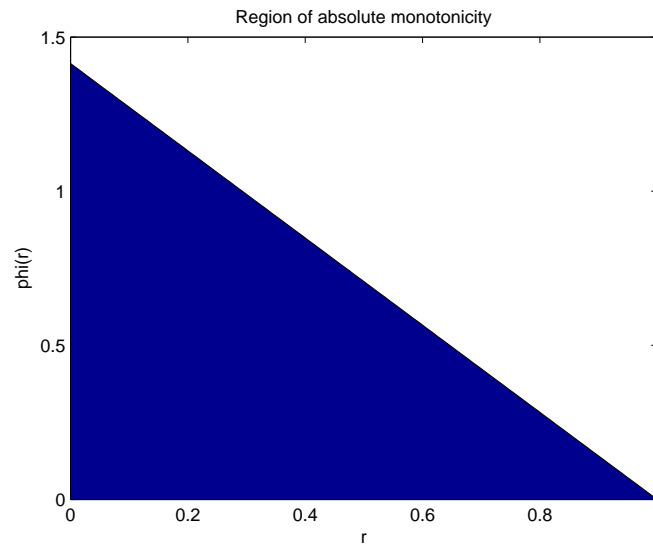


Figure 1: Region of absolute monotonicity for method (13).

An implementation taking into account the memory requirements could be realized using **six vectors in memory** in the following way (We do not claim that this is the best way to do it; note however that recombination of previously computed values is generally not possible due to the separate treatment of the two vector fields; for the same reason, implementation of a forward Euler predictor does not seem natural, so the value of the previous stage is used; the copy statement `yaux := F(ystage)` does not have to be performed in practice but is inserted for clarity of the presentation):

```
gamma := 1-1/sqrt(2);
ynought := yold;
yinc := yold;

% solve ystage = yold + h * gamma * G(ystage);      % first stage
do until convergence
    ystage := yinc + h * gamma * G(yold);
    yold := ystage;
end do

yaux := F(ystage);
ysum := ynought + 0.5 * h * yaux + 0.5 * (ystage - yinc) / gamma;
yinc := ynought + h * yaux + (1-2*gamma) * (ystage - yinc) / gamma;
```

Δt	error	order	C
2.0313e-02	2.1136e-03		
1.0156e-02	5.3037e-04	1.99	5.0162e+00
5.0781e-03	1.3289e-04	2.00	5.0662e+00
2.5391e-03	3.3263e-05	2.00	5.1061e+00
1.2695e-03	8.3209e-06	2.00	5.1319e+00
6.3477e-04	2.0808e-06	2.00	5.1481e+00
3.1738e-04	5.2020e-07	2.00	5.1649e+00
1.5869e-04	1.3002e-07	2.00	5.1762e+00
7.9346e-05	3.2576e-08	2.00	5.0248e+00
3.9673e-05	8.1186e-09	2.00	5.3999e+00

Table 1: Empirical convergence order for (13) applied to (10).

```

% solve ystage = yinc + h * gamma * G(ystage);      % second stage
do until convergence
  ystage := yinc + h * gamma * G(yold);
  yold := ystage;
end do

yaux := F(ystage);
ysum := ysum + 0.5 * h * yaux + 0.5 * (ystage - yinc) / gamma;

ynew := ysum;                                     % finish subinterval

```

The convergence order two is confirmed empirically in Table 1. The stability function (6) was computed by MAPLE to be

$$R(z) = 2 \frac{\sqrt{4 - 8 \Re(z) + 8 \Re(z) \sqrt{2} - 8 \Re(z)^2 \sqrt{2} + 12 \Re(z)^2 + \Im(z)^4 + 4 \Re(z) \sqrt{2} \Im(z)^2 + 12 \Re(z)^2 \Im(z)^2 - 4 \Re(z) \Im(z)^2 - 8 \Im(z)^2 \Re(z)^2 \sqrt{2}}}{(2 - 2 \Re(z) + \Re(z) \sqrt{2})^2}.$$

The stability region where $|R(z)| < 1$ seems to cover an unbounded region in the negative half-plane. A plot is given in Figure 2 in two different resolutions to demonstrate that the stability region is entirely located in the left half plane, tangent to the imaginary axis and appears to be unbounded as $\Re(z) \rightarrow -\infty$. Note that this means that the schemes are $A(\alpha)$ -stable with $\alpha = \frac{\pi}{2}$, but not A -stable [10]. Moreover, $\lim_{\Re(z) \rightarrow -\infty} R(z) = 0$.

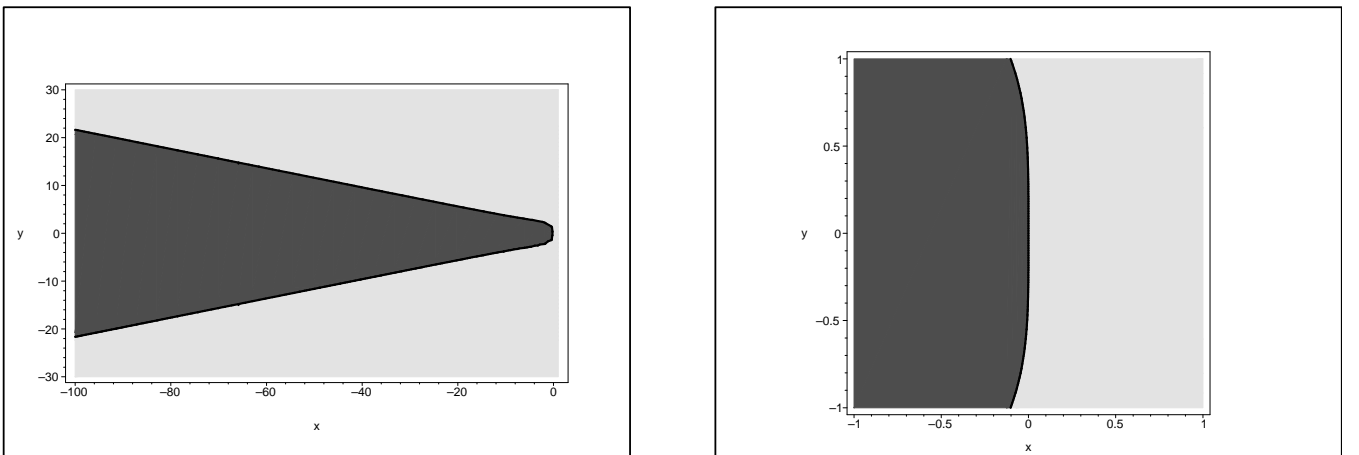


Figure 2: Stability region of method (13).

The stability function of the implicit scheme \tilde{A} is given by

$$R_{\tilde{A}}(z) = 2 \frac{(1 + \sqrt{2})(1 + z + \sqrt{2})}{(2 - z + \sqrt{2})^2}. \quad (14)$$

θ	$g(\theta)$
$-\pi$	$2 \frac{(1+\sqrt{2})(1-4\mu+\sqrt{2})}{(2+4\mu+\sqrt{2})^2}$
$-\frac{3\pi}{4}$	$-2 \frac{(1+\sqrt{2})(-1+2\mu+\mu\sqrt{2}-\sqrt{2})}{(2+2\mu+\mu\sqrt{2}+\sqrt{2})^2}$
$-\frac{\pi}{2}$	$2 \frac{(1+\sqrt{2})(1-2\mu+\sqrt{2})}{(2+2\mu+\sqrt{2})^2}$
$-\frac{\pi}{4}$	$2 \frac{(1+\sqrt{2})(1-2\mu+\mu\sqrt{2}+\sqrt{2})}{(-2-2\mu+\mu\sqrt{2}-\sqrt{2})^2}$
0	1
$\frac{\pi}{4}$	$2 \frac{(1+\sqrt{2})(1-2\mu+\mu\sqrt{2}+\sqrt{2})}{(-2-2\mu+\mu\sqrt{2}-\sqrt{2})^2}$
$\frac{\pi}{2}$	$2 \frac{(1+\sqrt{2})(1-2\mu+\sqrt{2})}{(2+2\mu+\sqrt{2})^2}$
$\frac{3\pi}{4}$	$-2 \frac{(1+\sqrt{2})(-1+2\mu+\mu\sqrt{2}-\sqrt{2})}{(2+2\mu+\mu\sqrt{2}+\sqrt{2})^2}$
π	$2 \frac{(1+\sqrt{2})(1-4\mu+\sqrt{2})}{(2+4\mu+\sqrt{2})^2}$

Table 2: Values of $g(\theta, \mu)$ for some θ , IMEX scheme (13), three point space discretization.

A plot of the related stability region can be seen in Figure 3. The scheme \tilde{A} appears to be A -stable and satisfies $\lim_{\Re(z) \rightarrow -\infty} R_{\tilde{A}}(z) = 0$. We thus conjecture that the scheme is L -stable.

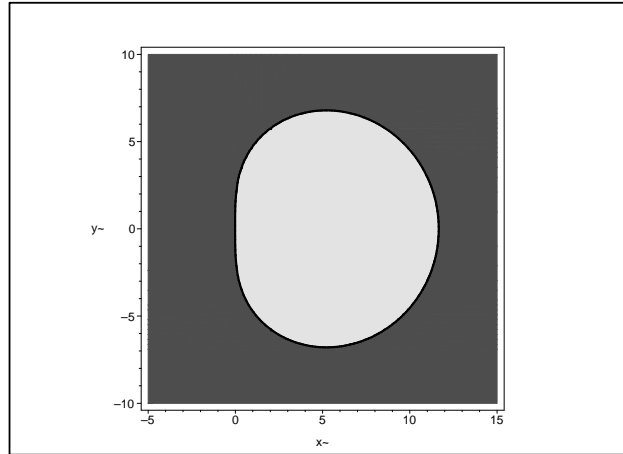


Figure 3: Stability region of the implicit method \tilde{A} in (13).

The dissipativity analysis for the implicit scheme defined by \tilde{A} yields the amplification factors (15) and (16) for the standard three-point space discretization (11) and the new fourth order stencil (12), respectively. The amplification factors are evaluated at the points $\theta \in \{-\pi + \ell\pi/4 : \ell = 0, \dots, 8\}$ in Tables 2 and 3, respectively.

$$g(\theta, \mu) = -\frac{(1 + \sqrt{2})(2\mu \cos(\theta) + 1 - 2\mu + \sqrt{2})}{-2\mu^2 (\cos(\theta))^2 + 4\mu \cos(\theta) + 4\mu^2 \cos(\theta) + 2\mu \cos(\theta) \sqrt{2} - 3 - 4\mu - 2\sqrt{2} - 2\mu^2 - 2\mu \sqrt{2}}. \quad (15)$$

The first positive zero of $g(\pi, \mu)$ is ≈ 0.603 for the three-point scheme (11), where the function changes its sign, and $|g(\pi, \mu)|$ never exceeds 1.

$$g(\theta, \mu) = \frac{6(1 + \sqrt{2})(-\mu(\cos(\theta))^2 - 7\mu + 8\mu \cos(\theta) + 3 + 3\sqrt{2})}{\mu^2 (\cos(\theta))^4 - 16\mu^2 (\cos(\theta))^3 + 78\mu^2 (\cos(\theta))^2 + (12 + 6\sqrt{2})\mu (\cos(\theta))^2 - 112\mu^2 \cos(\theta) - 48\mu \cos(\theta) \sqrt{2} - 96\mu \cos(\theta) + 54 + 49\mu^2 + 42\sqrt{2}\mu + 84\mu + 36\sqrt{2}}. \quad (16)$$

The first positive zero of $g(\pi, \mu)$ for the fourth order space discretization (12) is ≈ 0.452 , where the function changes its sign. The modulus never exceeds 1.

θ	$g(\theta)$
$-\pi$	$6 \frac{(1+\sqrt{2})(3-16\mu+3\sqrt{2})}{(6+16\mu+3\sqrt{2})^2}$
$-\frac{3\pi}{4}$	$-12 \frac{(1+\sqrt{2})(-6+15\mu+8\sqrt{2}\mu-6\sqrt{2})}{(12+15\mu+8\sqrt{2}\mu+6\sqrt{2})^2}$
$-\frac{\pi}{2}$	$6 \frac{(1+\sqrt{2})(3-7\mu+3\sqrt{2})}{(6+7\mu+3\sqrt{2})^2}$
$-\frac{\pi}{4}$	$12 \frac{(1+\sqrt{2})(6-15\mu+8\sqrt{2}\mu+6\sqrt{2})}{(-12-15\mu+8\sqrt{2}\mu-6\sqrt{2})^2}$
0	1
$\frac{\pi}{4}$	$12 \frac{(1+\sqrt{2})(6-15\mu+8\sqrt{2}\mu+6\sqrt{2})}{(-12-15\mu+8\sqrt{2}\mu-6\sqrt{2})^2}$
$\frac{\pi}{2}$	$6 \frac{(1+\sqrt{2})(3-7\mu+3\sqrt{2})}{(6+7\mu+3\sqrt{2})^2}$
$\frac{3\pi}{4}$	$-12 \frac{(1+\sqrt{2})(-6+15\mu+8\sqrt{2}\mu-6\sqrt{2})}{(12+15\mu+8\sqrt{2}\mu+6\sqrt{2})^2}$
π	$6 \frac{(1+\sqrt{2})(3-16\mu+3\sqrt{2})}{(6+16\mu+3\sqrt{2})^2}$

Table 3: Values of $g(\theta, \mu)$ for some θ , IMEX scheme (13), fourth order space discretization.

Modification of γ

We may conceive of optimizing the method (13) by adapting the value of the parameter γ in the definition of \tilde{A} according to the dissipativity properties. The region of absolute monotonicity depends on γ as follows [12]:

$$\mathcal{R}(\tilde{A}) = \begin{cases} \frac{1}{1-3\gamma}, & 0 \leq \gamma \leq \frac{1}{4}, \\ \frac{1-2\gamma}{2\gamma^2-4\gamma+1} - \sqrt{\frac{4\gamma-1}{(2\gamma^2-4\gamma+1)^2}}, & 1/4 < \gamma < 1 - \frac{1}{\sqrt{2}}, \\ 1 + \sqrt{2}, & \gamma = 1 - \frac{1}{\sqrt{2}}, \\ \frac{1-2\gamma}{2\gamma^2-4\gamma+1} + \sqrt{\frac{4\gamma-1}{(2\gamma^2-4\gamma+1)^2}}, & 1 - \frac{1}{\sqrt{2}} < \gamma \leq \frac{1}{2}, \end{cases} \quad (17)$$

$$\mathcal{R}(A, \tilde{A}) = \begin{cases} \left\{ 0 \leq r_1 \leq 1, 0 \leq r_2 \leq \frac{1-r_1}{1-\gamma} \right\}, & 0 \leq \gamma \leq \frac{1}{3}, \\ \left\{ 0 \leq r_1 \leq \frac{1-2\gamma}{\gamma}, 0 \leq r_2 \leq \frac{1-r_1}{1-\gamma} \right\}, & \frac{1}{3} \leq \gamma \leq \frac{1}{2}. \end{cases} \quad (18)$$

A plot of the function $\mathcal{R}(\tilde{A})$ in dependence of γ is given in Figure 4.

The regions of absolute monotonicity $\mathcal{R}(A, \tilde{A})$ for the values $\gamma \in \{0.1, 0.2, 0.3, 0.4, 0.45\}$ are plotted in Figure 5.

The stability regions for the IMEX schemes for the different values of γ cover bounded subdomains of the left half plane for $\gamma < 0.25$, while for $\gamma \geq 0.25$, the stability regions appear to cover an unbounded domain in the left half plane. However, even in the cases with unbounded stability regions, in general $\lim_{\Re(z) \rightarrow -\infty} R(z) \neq 0$. The boundaries of the stability regions are plotted in Figure 6, where values equal to 0 represent unbounded stability regions. The implicit schemes \tilde{A} show the same stability behaviour as concerns both the boundaries of the stability regions and the limits for $\Re(z) \rightarrow -\infty$.

With a MATLAB implementation it was demonstrated that the convergence order two is retained also for the modified values of γ . The error constant depends on γ , however. In Figure 7 we plot the error constant as a function of γ , where the error is determined at $t = 1.3$. We note that for small γ , the error constant decreases as γ grows, while for $\gamma > 0.183$ the constant grows monotonically. This behaviour does not appear to be related to the results we obtain below for the dissipativity analysis, where for $\gamma = 0.25$ the behaviour changes.

To assess the dissipativity properties of the modified scheme, in Figure 8 we vary γ and compute for 30 values of γ spaced equidistantly in the interval $[0.15, 1.1 - 1/\sqrt{2}]$ the first positive zero of $g(\pi, \mu)$ and the point μ where the modulus of this function exceeds 1. This analysis is given for the three point space discretization (11) in Figure 8. We observe that for $\gamma < 0.25$, the point where $|g(\pi, \mu)| > 1$ increases with increasing γ and in that parameter range g has no positive zero. For $\gamma > 0.25$, $|g(\pi, \mu)| < 1$ for all $\mu > 0$, but there is a positive root of that function where the solution might start oscillating. The picture is similar for the dissipative fourth order space discretization (12), where the critical values of μ scale with a factor of 0.75 as compared to the three point scheme.

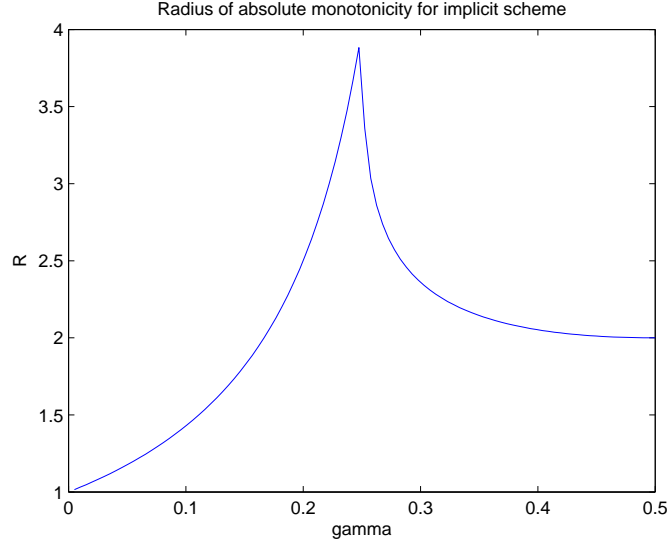


Figure 4: Radius of absolute monotonicity $\mathcal{R}(\tilde{A})$ as a function of γ for (13).

2.2 An SSP2(3,3,2) Method

[14] gives an IMEX SSP2(3,3,2) method with nontrivial region of absolute monotonicity:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline A & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \quad \begin{array}{c|ccc} \frac{1}{5} & \frac{1}{5} & 0 & 0 \\ \frac{3}{10} & \frac{1}{10} & \frac{1}{5} & 0 \\ 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \hline \tilde{A} & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{array} \quad (19)$$

This is a modification of a scheme from [25], where the latter turned out to have a trivial region of absolute monotonicity. It holds $\mathcal{R}(A) = 2$ and $R(\tilde{A}) = \frac{5}{9}(\sqrt{70} - 4)$, and

$$\mathcal{R}(A, \tilde{A}) = \{(r, \tilde{r}) : 0 \leq r \leq 2, 0 \leq \tilde{r} \leq \phi(r)\},$$

where

$$\phi(r) = \begin{cases} \frac{1}{4}(-28 + 9r) + \frac{1}{4}\sqrt{1264 - 984r + 201r^2} & \text{if } 0 \leq r \leq \frac{1}{64}(119 - \sqrt{721}), \\ \frac{5}{36}(-16 + r) + \frac{5\sqrt{7}}{36}\sqrt{160 - 128r + 31r^2} & \text{if } \frac{1}{64}(119 - \sqrt{721}) \leq r \leq 2. \end{cases}$$

A plot of $\mathcal{R}(A, \tilde{A})$ is given in Figure 9.

An implementation taking into account the memory requirements could be realized using **seven vectors in memory** in the following way (We do not claim that this is the best way to do it; note however that recombination of previously computed values is generally not possible due to the separate treatment of the two vector fields; for the same reason, implementation of a forward Euler predictor does not seem natural, so the value of the previous stage is used; the copy statements `yaux := F(ystage)` and `yaux := yaux + F(ystage)` do not have to be performed in practice but are inserted for clarity of the presentation):

```

ynought := yold;
yinc := yold;

% solve ystage = yold + 0.2 * h * G(ystage);      % first stage
do until convergence
  ystage := yinc + 0.2 * h * G(yold);
  yold := ystage;
end do

yaux := F(ystage);
ysum := ynought + 5/3 * (ystage - yinc);
yinc := ynought + 0.5 * h * yaux + 0.5 * (ystage - yinc);

```

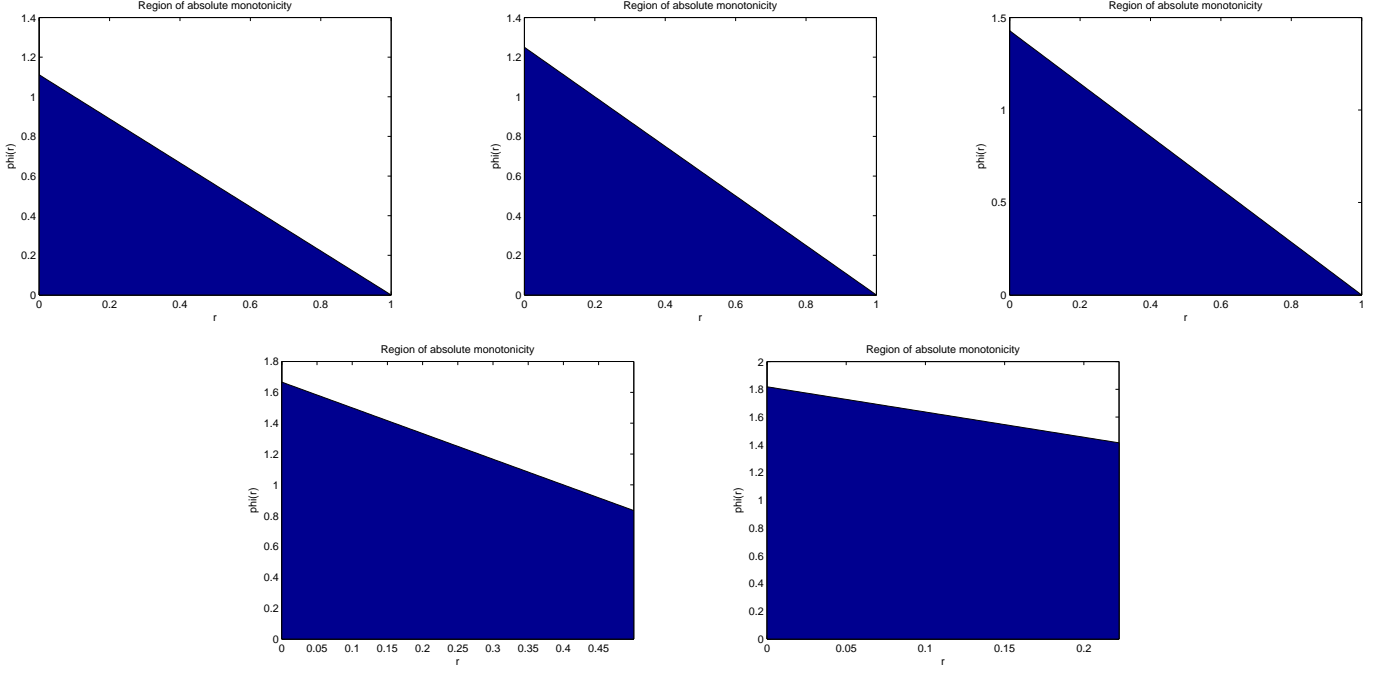


Figure 5: Regions of absolute monotonicity $\mathcal{R}(A, \tilde{A})$ for $\gamma \in \{0.1, 0.2, 0.3, 0.4, 0.45\}$ for (13).

```

% solve ystage = yinc + 0.2 * h * G(ystage);      % second stage
do until convergence
    ystage := yinc + 0.2 * h * G(yold);
    yold := ystage;
end do

yaux := yaux + F(ystage);
ysum := ysum + 5/3 * (ystage - yinc);
yinc := 0.5 * h * yaux + ysum;

% solve ystage = yinc + 1/3 * h * G(ystage);      % third stage
do until convergence
    ystage := yinc + 1/3 * h * G(yold);
    yold := ystage;
end do

ysum := ysum + 1/3 * h * yaux + (ystage - yinc) + h/3 * F(ystage);
ynew := ysum;                                     % finish subinterval

```

The convergence order two is confirmed empirically in Table 4.

The stability function (6) was computed by MAPLE. The stability region where $|R(z)| < 1$ seems to cover an unbounded sector in the negative half-plane. A plot is given in Figure 10 in two different resolutions to demonstrate that the stability region is entirely located in the left half plane, tangent to the imaginary axis and appears to be unbounded as $\Re(z) \rightarrow -\infty$. Moreover, $\lim_{\Re(z) \rightarrow -\infty} R(z) = 0$. The stability function of the implicit scheme is given by

$$R_{\tilde{A}}(z) = \frac{-150 - 40z + 9z^2}{2(-5+z)^2(-3+z)}. \quad (20)$$

A plot can be seen in Figure 11. Note that the stability region is not connected. The method is A -stable, however. Moreover, the scheme \tilde{A} satisfies $\lim_{\Re(z) \rightarrow -\infty} R_{\tilde{A}}(z) = 0$.

The dissipativity analysis for the implicit scheme defined by \tilde{A} yields the amplification factors for the standard three-point space discretization (11) and the new fourth order stencil (12), where the former is given in (21). The corresponding formula for (12)

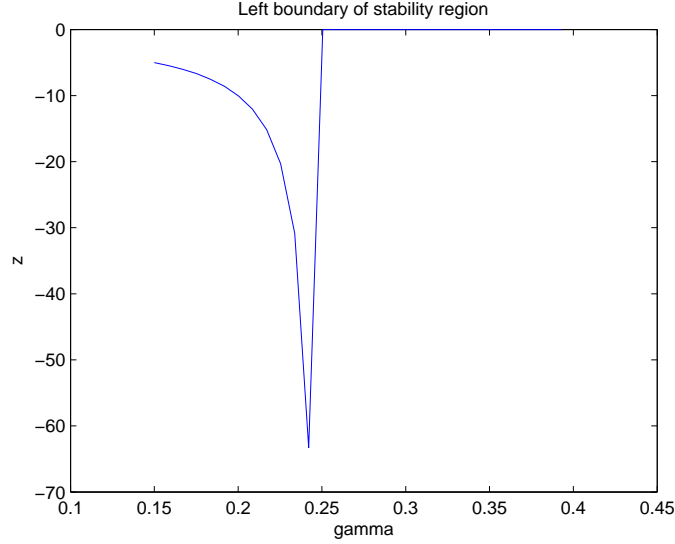


Figure 6: Left boundary of the stability region as a function of γ for (13).

Δt	error	order	C
2.0313e-02	3.3570e-03		
1.0156e-02	8.3585e-04	2.01	8.3231e+00
5.0781e-03	2.0867e-04	2.00	8.1801e+00
2.5391e-03	5.2137e-05	2.00	8.1267e+00
1.2695e-03	1.3031e-05	2.00	8.1043e+00
6.3477e-04	3.2573e-06	2.00	8.0963e+00
3.1738e-04	8.1420e-07	2.00	8.0974e+00
1.5869e-04	2.0362e-07	2.00	8.0529e+00
7.9346e-05	5.0890e-08	2.00	8.1128e+00
3.9673e-05	1.2669e-08	2.01	8.5656e+00

Table 4: Empirical convergence order for (19) applied to (10).

is not given for layout reasons. The amplification factors are evaluated at the points $\theta \in \{-\pi + \ell\pi/4 : \ell = 0, \dots, 8\}$ in Tables 5 and 6, respectively.

$$g(\theta, \mu) = \frac{18\mu^2 (\cos(\theta))^2 + 18\mu^2 - 40\mu \cos(\theta) - 36 \cos(\theta) \mu^2 - 75 + 40\mu}{8\mu^3 (\cos(\theta))^3 - 52\mu^2 (\cos(\theta))^2 - 24\mu^3 (\cos(\theta))^2 + 110\mu \cos(\theta) + 104 \cos(\theta) \mu^2 + 24\mu^3 \cos(\theta) - 75 - 110\mu - 52\mu^2 - 8\mu^3}. \quad (21)$$

The first positive zero of $g(\pi, \mu)$ is ≈ 0.606 for the three-point space discretization, where the function changes its sign, and $|g(\pi, \mu)|$ never exceeds 1.

The first positive zero of $g(\pi, \mu)$ is ≈ 0.455 for the fourth order space discretization, where the function changes its sign, and $|g(\pi, \mu)|$ never exceeds 1.

2.2.1 The Explicit SSP(3,2) Scheme

Finally, we give the corresponding results for the explicit tableau A from (19), as this was found to excel in its practical value over the classical explicit schemes [29]. This scheme was first published in [20] and later denoted by $SSP(3, 2)$ and declared the optimal second order scheme with three stages in [30] as well as in [26], where it was denoted by $SSPRK(3, 2)$, and independently also in [7]. An implementation for the generic equation $\dot{y} = F(y)$, $y(0) = y_0$ taking into account the memory requirements could be realized using **two vectors in memory** in the following way:

```
ynought := yold;
ystage := ynought + 0.5 * h * F(ynought);
ystage := ystage + 0.5 * h * F(ystage);
```

θ	$g(\theta)$
$-\pi$	$-\frac{-75+80\mu+72\mu^2}{(5+4\mu)^2(3+4\mu)}$
$-\frac{3\pi}{4}$	$-\frac{-75+20\mu\sqrt{2}+40\mu+27\mu^2+18\mu^2\sqrt{2}}{(5+\mu\sqrt{2}+2\mu)^2(3+\mu\sqrt{2}+2\mu)}$
$-\frac{\pi}{2}$	$-\frac{-75+18\mu^2+40\mu}{(5+2\mu)^2(3+2\mu)}$
$-\frac{\pi}{4}$	$-\frac{75+20\mu\sqrt{2}-40\mu-27\mu^2+18\mu^2\sqrt{2}}{(-5+\mu\sqrt{2}-2\mu)^2(-3+\mu\sqrt{2}-2\mu)}$
0	1
$\frac{\pi}{4}$	$-\frac{75+20\mu\sqrt{2}-40\mu-27\mu^2+18\mu^2\sqrt{2}}{(-5+\mu\sqrt{2}-2\mu)^2(-3+\mu\sqrt{2}-2\mu)}$
$\frac{\pi}{2}$	$-\frac{-75+18\mu^2+40\mu}{(5+2\mu)^2(3+2\mu)}$
$\frac{3\pi}{4}$	$-\frac{-75+20\mu\sqrt{2}+40\mu+27\mu^2+18\mu^2\sqrt{2}}{(5+\mu\sqrt{2}+2\mu)^2(3+\mu\sqrt{2}+2\mu)}$
π	$-\frac{-75+80\mu+72\mu^2}{(5+4\mu)^2(3+4\mu)}$

Table 5: Values of $g(\theta, \mu)$ for some θ , IMEX scheme (19), three point space discretization.

θ	$g(\theta)$
$-\pi$	$-9\frac{320\mu+384\mu^2-225}{(15+16\mu)^2(9+16\mu)}$
$-\frac{3\pi}{4}$	$-9\frac{-1800+1200\mu+640\mu\sqrt{2}+1059\mu^2+720\mu^2\sqrt{2}}{(30+15\mu+8\mu\sqrt{2})^2(18+15\mu+8\mu\sqrt{2})}$
$-\frac{\pi}{2}$	$-9/2\frac{-450+280\mu+147\mu^2}{(7\mu+15)^2(7\mu+9)}$
$-\frac{\pi}{4}$	$-9\frac{1800-1200\mu+640\mu\sqrt{2}-1059\mu^2+720\mu^2\sqrt{2}}{(-30-15\mu+8\mu\sqrt{2})^2(-18-15\mu+8\mu\sqrt{2})}$
0	1
$\frac{\pi}{4}$	$-9\frac{1800-1200\mu+640\mu\sqrt{2}-1059\mu^2+720\mu^2\sqrt{2}}{(-30-15\mu+8\mu\sqrt{2})^2(-18-15\mu+8\mu\sqrt{2})}$
$\frac{\pi}{2}$	$-9/2\frac{-450+280\mu+147\mu^2}{(7\mu+15)^2(7\mu+9)}$
$\frac{3\pi}{4}$	$-9\frac{-1800+1200\mu+640\mu\sqrt{2}+1059\mu^2+720\mu^2\sqrt{2}}{(30+15\mu+8\mu\sqrt{2})^2(18+15\mu+8\mu\sqrt{2})}$
π	$-9\frac{320\mu+384\mu^2-225}{(15+16\mu)^2(9+16\mu)}$

Table 6: Values of $g(\theta, \mu)$ for some θ , IMEX scheme (19), fourth order space discretization.

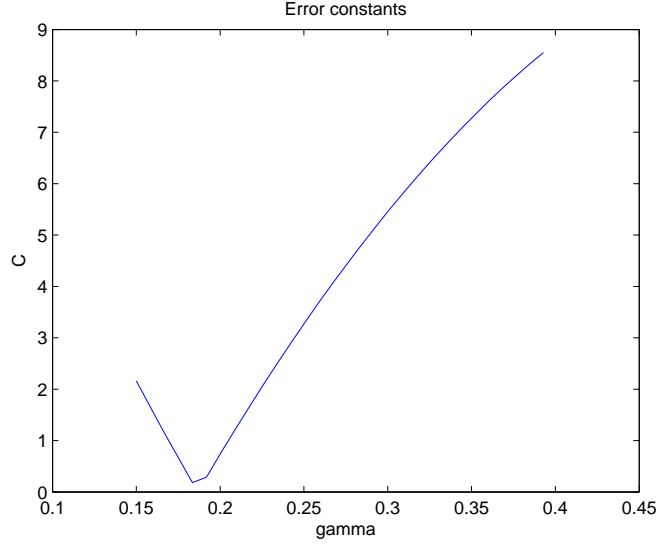


Figure 7: Error constant as a function of γ for (13).



Figure 8: Dissipativity analysis for (13) for different values of γ , three point space discretization (left) and fourth order discretization (right).

`ynew := ynought + (2/3) * (ystage - ynought) + (h/3) * F(ystage);`

This is equivalent to coding the scheme according to the α - β (Shu-Osher) representation [29] given by [30]. The convergence order two is confirmed empirically in Table 7. The stability function was computed by MAPLE to equal

$$R_A(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{12}.$$

The stability region where $|R(z)| < 1$ occupies a bounded region in the negative half-plane, and is tangent to the imaginary axis, see Figure 12. Note that $\lim_{\Re(z) \rightarrow -\infty} |R(z)| = \infty$.

The point where the stability region intersects the negative real half-line is located at $x \approx -4.519$.

The dissipativity analysis for A yields the amplification factor (22) for the standard three-point space discretization (11). The corresponding formula for the new fourth order stencil (12) is not given for layout reasons. The amplification factors are evaluated at the points $\theta \in \{-\pi + \ell\pi/4 : \ell = 0, \dots, 8\}$ in Tables 8 and 9, respectively.

$$g(\mu, \theta) = -2\mu^3 (\cos(\theta))^2 - 2/3\mu^3 + 2\mu^2 (\cos(\theta))^2 + 2\mu^2 + 2/3\mu^3 (\cos(\theta))^3 + 2\mu^3 \cos(\theta) - 4 \cos(\theta) \mu^2 + 2 \cos(\theta) \mu + 1 - 2\mu \quad (22)$$

For the three-point space discretization (11), the first positive zero of $g(\pi, \mu)$ is about 0.896, where the function changes its sign, and $g(\pi, \mu) = -1$ for $\mu \approx 1.129$.

Table 9 gives the dissipativity analysis for A from (19) for the fourth order space discretization (12).

The first positive zero of $g(\pi, \mu)$ is about 0.672 for the fourth-order space discretization (12), where the function changes its sign, and $g(\pi, \mu) = -1$ at $\mu \approx 0.847$.

Δt	error	order	C
6.5000e-01	6.7065e-01	—	—
3.2500e-01	3.2895e-01	1.03	1.0441e+00
1.6250e-01	1.2210e-01	1.43	1.6407e+00
8.1250e-02	3.7034e-02	1.72	2.7860e+00
4.0625e-02	1.0074e-02	1.88	4.1317e+00
2.0313e-02	2.6117e-03	1.95	5.1606e+00
1.0156e-02	6.6362e-04	1.98	5.7782e+00
5.0781e-03	1.6716e-04	1.99	6.1192e+00
2.5391e-03	4.1943e-05	1.99	6.3051e+00
1.2695e-03	1.0505e-05	2.00	6.4068e+00

Table 7: Empirical convergence order for A in (19) applied to (10).

θ	$g(\theta)$
$-\pi$	$1 - 4\mu + 8\mu^2 - 16/3\mu^3$
$-\frac{3\pi}{4}$	$1 - \mu\sqrt{2} - 2\mu + 3\mu^2 + 2\mu^2\sqrt{2} - 7/6\mu^3\sqrt{2} - 5/3\mu^3$
$-\frac{\pi}{2}$	$1 - 2\mu + 2\mu^2 - 2/3\mu^3$
$-\frac{\pi}{4}$	$1 + \mu\sqrt{2} - 2\mu + 3\mu^2 - 2\mu^2\sqrt{2} + 7/6\mu^3\sqrt{2} - 5/3\mu^3$
0	1
$\frac{\pi}{4}$	$1 + \mu\sqrt{2} - 2\mu + 3\mu^2 - 2\mu^2\sqrt{2} + 7/6\mu^3\sqrt{2} - 5/3\mu^3$
$\frac{\pi}{2}$	$1 - 2\mu + 2\mu^2 - 2/3\mu^3$
$\frac{3\pi}{4}$	$1 - \mu\sqrt{2} - 2\mu + 3\mu^2 + 2\mu^2\sqrt{2} - 7/6\mu^3\sqrt{2} - 5/3\mu^3$
π	$1 - 4\mu + 8\mu^2 - 16/3\mu^3$

Table 8: Values of $g(\theta, \mu)$ for some θ , explicit scheme A from (19), three point space discretization.

θ	$g(\theta)$
$-\pi$	$1 - 16/3\mu + \frac{128}{9}\mu^2 - \frac{1024}{81}\mu^3$
$-\frac{3\pi}{4}$	$1 - 5/2\mu - 4/3\mu\sqrt{2} + \frac{353}{72}\mu^2 + 10/3\mu^2\sqrt{2} - \frac{1015}{288}\mu^3 - \frac{803}{324}\mu^3\sqrt{2}$
$-\frac{\pi}{2}$	$1 - 7/3\mu + \frac{49}{18}\mu^2 - \frac{343}{324}\mu^3$
$-\frac{\pi}{4}$	$1 - 5/2\mu + 4/3\mu\sqrt{2} + \frac{353}{72}\mu^2 - 10/3\mu^2\sqrt{2} - \frac{1015}{288}\mu^3 + \frac{803}{324}\mu^3\sqrt{2}$
0	1
$\frac{\pi}{4}$	$1 - 5/2\mu + 4/3\mu\sqrt{2} + \frac{353}{72}\mu^2 - 10/3\mu^2\sqrt{2} - \frac{1015}{288}\mu^3 + \frac{803}{324}\mu^3\sqrt{2}$
$\frac{\pi}{2}$	$1 - 7/3\mu + \frac{49}{18}\mu^2 - \frac{343}{324}\mu^3$
$\frac{3\pi}{4}$	$1 - 5/2\mu - 4/3\mu\sqrt{2} + \frac{353}{72}\mu^2 + 10/3\mu^2\sqrt{2} - \frac{1015}{288}\mu^3 - \frac{803}{324}\mu^3\sqrt{2}$
π	$1 - 16/3\mu + \frac{128}{9}\mu^2 - \frac{1024}{81}\mu^3$

Table 9: Values of $g(\theta, \mu)$ for some θ , explicit scheme A from (19), fourth order space discretization.

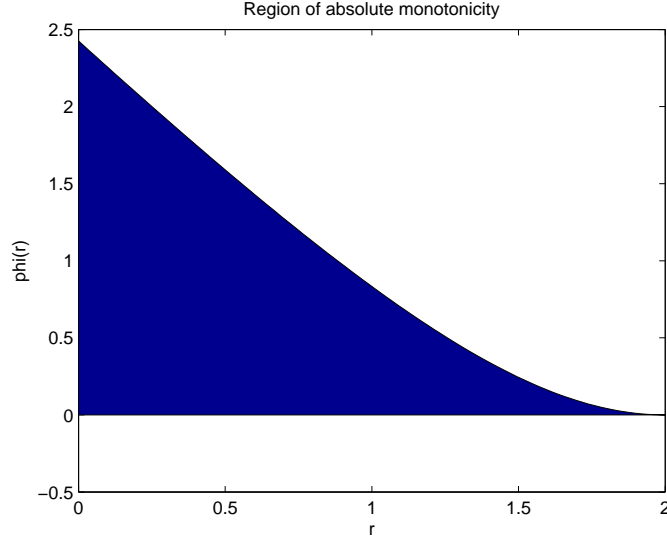


Figure 9: Region of absolute monotonicity for method (19).

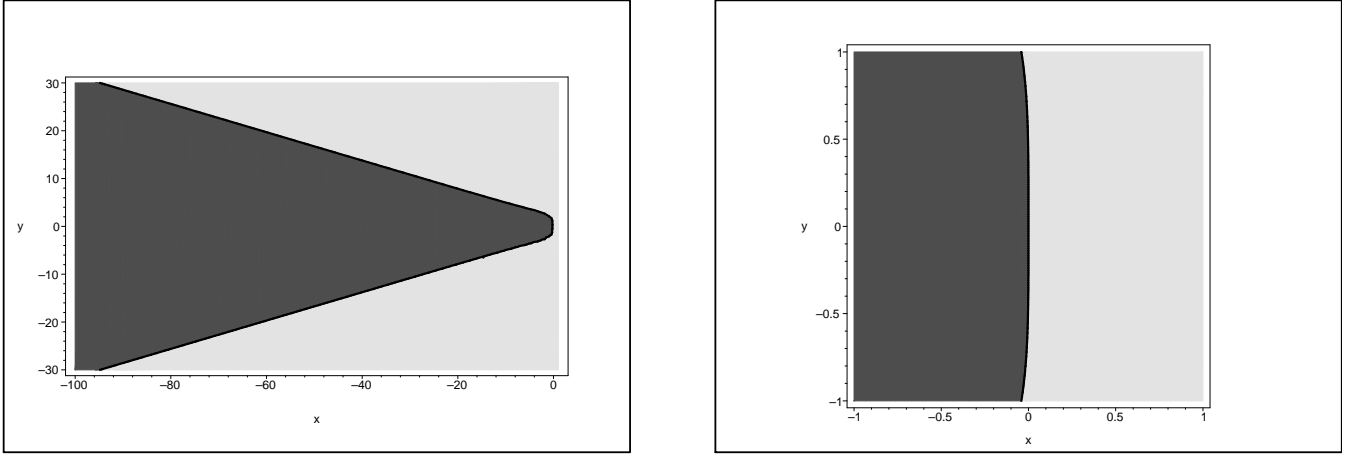


Figure 10: Stability region of method (19).

2.3 An SSP3(3,3,3) Method

[15] gives the following SSP3(3,3,3) method with nontrivial region of absolute monotonicity:

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} & 0 \\ \hline A & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{array} \quad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ 1 & \frac{14}{15} & \frac{1}{15} & 0 \\ \frac{1}{2} & \frac{7}{30} & \frac{1}{5} & \frac{1}{15} \\ \hline \tilde{A} & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{array} \quad (23)$$

This scheme requires a negative time step between the stages which is executed explicitly only in a predictor step (if implemented). It holds $\mathcal{R}(A) = 1$ and $\mathcal{R}(\tilde{A}) = \frac{5}{47}(13 - 2\sqrt{7})$, and

$$\mathcal{R}(A, \tilde{A}) = \{(r, \tilde{r}) : 0 \leq r \leq 1, 0 \leq \tilde{r} \leq \phi(r)\},$$

where

$$\phi(r) = \frac{15}{302} \left(28 - 25r - \sqrt{180 - 192r + 21r^2} \right).$$

A plot of $\mathcal{R}(A, \tilde{A})$ is given in Figure 13.

An implementation taking into account the memory requirements could be realized using **seven vectors in memory** in the following way (We do not claim that this is the best way to do it; note however that recombination of previously computed values

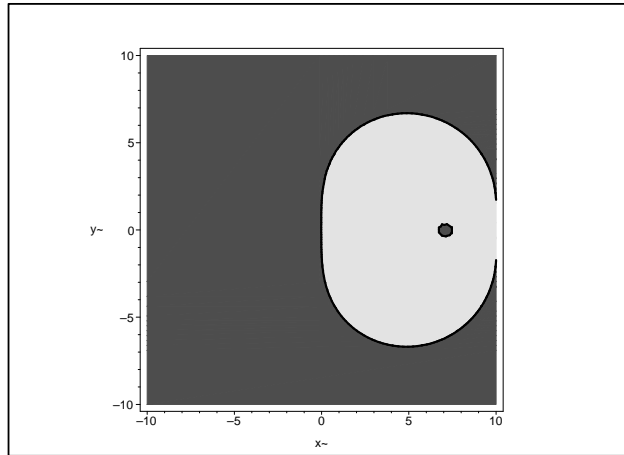


Figure 11: Stability region of the implicit method \tilde{A} in (19).

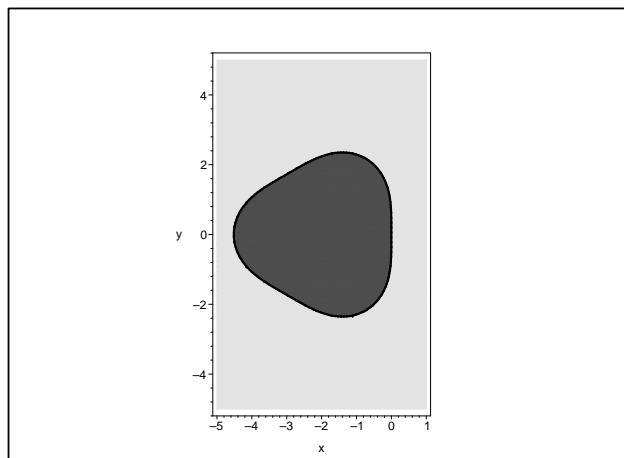


Figure 12: Stability region of method A from (19).

is generally not possible due to the separate treatment of the two vector fields; for the same reason, implementation of a forward Euler predictor does not seem natural, so the value of the previous stage is used; the copy statements `yaux := F(ystage)` and `yaux := yaux + F(ystage)` do not have to be performed in practice but are inserted for clarity of the presentation):

```

ynought := yold;
yinc1 := h * ( F(ynought) + (14/15)*G(ynought) ); % first stage corresponds with initial values
yinc2 := 0.25 * h * F(ynought) + (7/30) * h* G(ynought);
ysum := ynought + (h/6) * ( F(ynought) + G(ynought) );

% solve second stage
do until convergence
    ystage := ynought + yinc1 + (h/15) * G(yold);
    yold := ystage;
end do

yaux := F(ystage);
ysum := ysum + (h/6) * yaux + (15/6) * (ystage - ynought - yinc1);
yinc2 := yinc2 + 0.25 * h * yaux + 3 * (ystage - ynought - yinc1);

% solve third stage
do until convergence
    ystage := ynought + yinc2 + (h/15) * G(yold);
    yold := ystage;

```

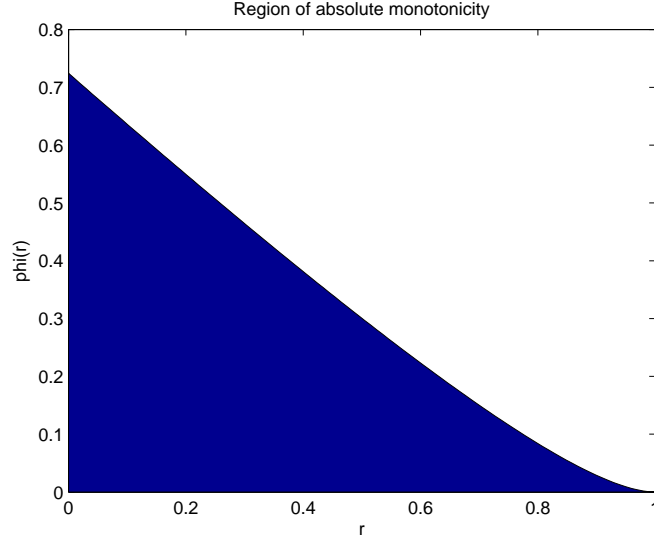


Figure 13: Region of absolute monotonicity for method (23).

Δt	error	order	C
6.5000e-01	3.6607e-01		
3.2500e-01	1.2504e-01	1.55	7.1368e-01
1.6250e-01	2.8093e-02	2.15	1.4076e+00
8.1250e-02	4.7746e-03	2.56	2.9260e+00
4.0625e-02	6.9320e-04	2.78	5.1762e+00
2.0313e-02	9.3123e-05	2.90	7.4111e+00
1.0156e-02	1.2056e-05	2.95	9.1233e+00
5.0781e-03	1.5332e-06	2.98	1.0265e+01
2.5391e-03	1.9328e-07	2.99	1.0977e+01
1.2695e-03	2.4209e-08	3.00	1.1605e+01

Table 10: Empirical convergence order for (23) applied to (10).

```

end do

yaux := F(ystage);
ysum := ysum + (2/3) * h * yaux + 10 * ( ystage - ynought - yinc2 );

ynew := ysum;                                     % finish subinterval

```

The convergence order three is confirmed empirically in Table 10. The stability function (6) was computed by MAPLE. The stability region where $|R(z)| < 1$ occupies a bounded region in the negative half-plane, and slightly overlaps the imaginary axis, see Figure 14. Note that $\lim_{\Re(z) \rightarrow -\infty} |R(z)| = \infty$.

The stability function of the implicit scheme is given by

$$R_{\tilde{A}}(z) = 1/2 \frac{450 + 390z + 167z^2 + 47z^3}{(-15 + z)^2}. \quad (24)$$

A plot can be seen in Figure 14. The point where the stability region intersects the negative real half-line is located at $x \approx -3.248$ for the IMEX scheme. The same value is computed for the stability region of the implicit scheme. However, the stability region of \tilde{A} extends further along the imaginary axis. Finally, $\lim_{\Re(z) \rightarrow -\infty} |R_{\tilde{A}}(z)| = \infty$.

The dissipativity analysis for the implicit scheme defined by \tilde{A} yields the amplification factor (25) for the standard three-point space discretization (11). The corresponding formula for the new fourth order stencil (12) is not given for layout reasons. The amplification factors are evaluated at the points $\theta \in \{-\pi + \ell\pi/4 : \ell = 0, \dots, 8\}$ in Tables 11 and 12, respectively.

$$g(\mu, \theta) = \frac{-564\mu^3 (\cos(\theta))^2 - 188\mu^3 + 334\mu^2 (\cos(\theta))^2 + 334\mu^2 + 188\mu^3 (\cos(\theta))^3 + 564\mu^3 \cos(\theta) - 668 \cos(\theta) \mu^2 + 390 \mu \cos(\theta) + 225 - 390 \mu}{4\mu^2 (\cos(\theta))^2 - 60 \mu \cos(\theta) - 8 \cos(\theta) \mu^2 + 225 + 60 \mu + 4 \mu^2} \quad (25)$$

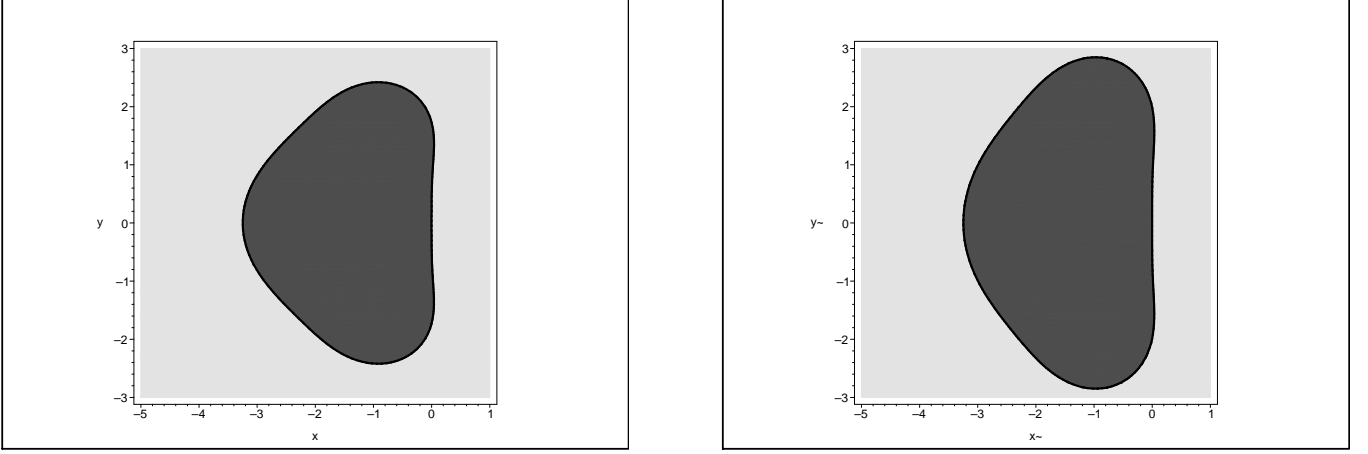


Figure 14: Stability regions of method (23) (left) and \tilde{A} .

θ	$g(\theta)$
$-\pi$	$-\frac{1504\mu^3+780\mu-1336\mu^2-225}{(15+4\mu)^2}$
$-\frac{3\pi}{4}$	$-\frac{-225+195\mu\sqrt{2}+390\mu-501\mu^2-334\mu^2\sqrt{2}+329\mu^3\sqrt{2}+470\mu^3}{(15+\mu\sqrt{2}+2\mu)^2}$
$-\frac{\pi}{2}$	$-\frac{-225+390\mu-334\mu^2+188\mu^3}{(15+2\mu)^2}$
$-\frac{\pi}{4}$	$\frac{225+195\mu\sqrt{2}-390\mu+501\mu^2-334\mu^2\sqrt{2}+329\mu^3\sqrt{2}-470\mu^3}{(-15+\mu\sqrt{2}-2\mu)^2}$
0	1
$\frac{\pi}{4}$	$\frac{225+195\mu\sqrt{2}-390\mu+501\mu^2-334\mu^2\sqrt{2}+329\mu^3\sqrt{2}-470\mu^3}{(-15+\mu\sqrt{2}-2\mu)^2}$
$\frac{\pi}{2}$	$-\frac{-225+390\mu-334\mu^2+188\mu^3}{(15+2\mu)^2}$
$\frac{3\pi}{4}$	$-\frac{-225+195\mu\sqrt{2}+390\mu-501\mu^2-334\mu^2\sqrt{2}+329\mu^3\sqrt{2}+470\mu^3}{(15+\mu\sqrt{2}+2\mu)^2}$
π	$-\frac{1504\mu^3+780\mu-1336\mu^2-225}{(15+4\mu)^2}$

Table 11: Values of $g(\theta, \mu)$ for some θ , IMEX scheme (23), three point space discretization.

For the three-point space discretization (11), the first positive zero of $g(\pi, \mu)$ is about 0.465, where the function changes its sign, and $g(\pi, \mu) = -1$ for $\mu \approx 0.812$.

The first positive zero of $g(\pi, \mu)$ is about 0.348 for the fourth-order space discretization (12), where the function changes its sign, and $g(\pi, \mu) = -1$ at $\mu \approx 0.609$.

3 The ANTARES code for astrophysical simulations

In this section, we describe the realization of IMEX methods in simulation routines for astrophysical models implemented in the ANTARES code [24]. The physical system which is simulated is the *fully compressible Navier–Stokes equation* which describes momentum conservation:

$$(\rho\mathbf{u})' + \nabla \cdot (\rho\mathbf{u} \otimes \mathbf{u} + p\mathbf{I}) = \rho\mathbf{g} + \nabla \cdot \tau. \quad (26)$$

The state variables in the model equations generally depend on the spatial variables (x, y, z) and time t . The (explicit) dependencies are stated in Table 13. For simplicity, we omit the dependencies in the problem specification (26)–(30). The model is completed by the *continuity equation*

$$\rho' + \nabla \cdot (\rho\mathbf{u}) = 0, \quad (27)$$

which ensures conservation of mass, and the *total energy equation*

$$e' + \nabla \cdot (\mathbf{u}(e + p)) = \rho(\mathbf{g} \cdot \mathbf{u}) + \nabla \cdot (\mathbf{u} \cdot \tau) + Q_{\text{rad}}, \quad (28)$$

θ	$g(\theta)$
$-\pi$	$-1/3 \frac{-6075+28080\mu-64128\mu^2+96256\mu^3}{(45+16\mu)^2}$
$-\frac{3\pi}{4}$	$-1/12 \frac{-97200+210600\mu+112320\mu\sqrt{2}-353706\mu^2-240480\mu^2\sqrt{2}+429345\mu^3+301928\mu^3\sqrt{2}}{(90+15\mu+8\mu\sqrt{2})^2}$
$-\frac{\pi}{2}$	$-1/6 \frac{-12150+24570\mu-24549\mu^2+16121\mu^3}{(45+7\mu)^2}$
$-\frac{\pi}{4}$	$1/12 \frac{97200-210600\mu+112320\mu\sqrt{2}+353706\mu^2-240480\mu^2\sqrt{2}-429345\mu^3+301928\mu^3\sqrt{2}}{(-90-15\mu+8\mu\sqrt{2})^2}$
0	1
$\frac{\pi}{4}$	$1/12 \frac{97200-210600\mu+112320\mu\sqrt{2}+353706\mu^2-240480\mu^2\sqrt{2}-429345\mu^3+301928\mu^3\sqrt{2}}{(-90-15\mu+8\mu\sqrt{2})^2}$
$\frac{\pi}{2}$	$-1/6 \frac{-12150+24570\mu-24549\mu^2+16121\mu^3}{(45+7\mu)^2}$
$\frac{3\pi}{4}$	$-1/12 \frac{-97200+210600\mu+112320\mu\sqrt{2}-353706\mu^2-240480\mu^2\sqrt{2}+429345\mu^3+301928\mu^3\sqrt{2}}{(90+15\mu+8\mu\sqrt{2})^2}$
π	$-1/3 \frac{-6075+28080\mu-64128\mu^2+96256\mu^3}{(45+16\mu)^2}$

Table 12: Values of $g(\theta, \mu)$ for some θ , IMEX scheme (23), fourth order space discretization.

which describes conservation of the latter. In the case of a two-component fluid, the system is augmented by the concentration equation of the second species,

$$(c\rho)' + \nabla \cdot (c\rho\mathbf{u}) = \nabla \cdot (\rho\kappa_c\nabla c). \quad (29)$$

The variables and parameters which appear in the model formulation are collected in Table 13.

$\rho = \rho(x, y, z, t)$	gas density
$c = c(x, y, z, t)$	concentration of second species
$\mathbf{u} = \mathbf{u}(x, y, z, t) = (u, v, w)^T$	flow velocity
$\rho\mathbf{u}$	momentum density
$\mathbf{u} \otimes \mathbf{u}$	Kronecker product
$p = p(T, \rho)$	gas pressure
$\mathbf{g} = (g, 0, 0)^T$	gravitational acceleration
$\tau = \tau(x, y, z, t)$	viscous stress tensor for zero bulk viscosity
μ	dynamic viscosity (appears in the definition of τ)
$e = e(x, y, z, t) = e_{\text{int}} + e_{\text{kin}}$	total energy density, the sum of internal and kinetic energy densities
$T = T(x, y, z, t)$	temperature
$Q_{\text{rad}} = Q_{\text{rad}}(x, y, z, t)$	radiative source term
$\chi_\nu = \chi_\nu(T, \rho)$	(specific) opacity at frequency ν
$\kappa = \kappa(T, \rho)$	radiative conductivity
$\kappa_c = \kappa(c, \rho)$	diffusion coefficient for species c
$I_\nu = I_\nu(\mathbf{r}), \mathbf{r} = \mathbf{r}(x, y, z)$	specific intensity along the ray of direction \mathbf{r}
$S_\nu = S_\nu(x, y, z)$	source function

Table 13: Variables and parameters in the equations (26)–(30).

The radiative source term Q_{rad} is determined as the stationary limit of the *radiative transfer equation*

$$\mathbf{r} \cdot \nabla I_\nu = \rho\chi_\nu(S_\nu - I_\nu), \quad (30)$$

which is solved for all ray directions \mathbf{r} and for all frequencies ν , resulting in the specific intensity $I_\nu(\mathbf{r})$, for details see [32]. S_ν here denotes the *source function*.

The equations of hydrodynamics (26), (27) and (28) are closed by the equation of state which describes the relation between the thermodynamic quantities. For the particular choice see [24].

For the initial condition, a slightly perturbed static model atmosphere or envelope is used which is equipped with a small seed velocity field or density perturbation to start dynamics away from equilibrium.

Boundary conditions are based on the assumption that all quantities are periodic in both horizontal directions. For the hydrodynamical equations, closed boundary conditions at the upper and lower boundary of the computational domain are used, but a recent development is to replace these by open boundary conditions. For the radiative transfer equation (30), incoming radiation at the boundary of the computational domain must be specified.

The ANTARES code (**A** Numerical **T**ool for **A**strophysical **R**ESearch) described in this report solves this system of equations numerically in either one, two, or three spatial dimensions on a rectangular grid (spherical coordinates with a logarithmically rectangular grid are also possible, i.e., the grid may be locally rectangular with logarithmic grading in the radial component). ANTARES allows the definition of several grids which can be nested inside each other to improve resolution in regions of interest. At the moment, ANTARES provides up to three levels of nested grids.

For the spatial discretization of the hyperbolic terms, discretizations of ENO (*essentially non-oscillatory* [28]) type are implemented. These comprise classical ENO methods, WENO (*weighted essentially non-oscillatory*) methods [28] (optionally in conjunction with Marquina flux splitting [3]) and CNO (*convex non-oscillatory*) schemes [22]. Each of the methods uses adaptive stencils which are chosen such as to avoid spurious oscillations in the computed solution. The spatial derivatives are calculated for each direction separately.

The parabolic terms are discretized by dissipative finite difference schemes [18] of fourth order. The *radiative heating rate* is determined by the *short characteristics method*, or by means of a diffusion approximation, where appropriate, while all other source terms are evaluated at the cell centers. For the time integration, *total variation diminishing* Runge–Kutta methods [27, 29] are employed.

ANTARES implements two different parallelization concepts. For architectures with distributed memory, domain decomposition is used and realized by an MPI implementation. In this approach each grid is split along the horizontal direction(s) and optionally, also along vertical ones, into subdomains. The memory required to store the computational variables for each subdomain is provided by the resources available to the dedicated CPU core which performs the computations necessary for that subdomain. In this way, each CPU core is mapped to a specific geometrical volume. However, since some supercomputers offer only a limited amount of memory per CPU core and because the domain decomposition approach just mentioned is not very efficient on small grids, ANTARES offers a second type of parallelization which can be used along with or independently of the former. It is based on a shared memory concept for each subdomain and is implemented through OpenMP directives. Thus, on each MPI node or on the entire grid, if no domain decomposition is performed, the most time consuming operations which can also be performed independently of each other are identified and computed in parallel. This approach scales only to a moderate number of CPU cores (in principle up to a few dozens), but allows improvement of the scaling and the computational speed of the domain decomposition based parallelization for a larger number of problems and for a greater variety of computing architectures.

3.1 Simulation of Semiconvection

In astrophysics, models of stellar structure predict settings where the heavier product of nuclear fusion provides stability to a zone which otherwise would be unstable to convective overturning. Such a zone would become convective if its composition were mixed. The question whether such a zone should be treated as if it were mixed or not is referred to as the *semiconvection problem*.

A thorough analysis of the semiconvection problem in two dimensions is given in [33]. In the following, we perform numerical simulations of semiconvection to demonstrate the efficiency of the IMEX methods as compared to the classical total variation diminishing Runge–Kutta time integrators.

The fluid is described by the multispecies Navier–Stokes equations presented in Section 3. Additionally, dimensionless quantities such as the Prandtl number Pr , the Lewis number Le , the Rayleigh number Ra and the stability parameter R_ρ are defined to determine the diffusivities κ , κ_c and the viscosity μ . The former quantities arise in the definition of the starting model but do not appear in the evolution equations (31).

Since intermolecular forces are neglected in this model, the equation of state of the fluid is assumed to be that of an ideal gas. The radiative source term Q_{rad} is modelled using the diffusion approximation which is adequate for stars where semiconvection zones occur close to the region of nuclear energy production in their deep interior (see [17, 33], Chap. 13.3 and 13-A in [32]).

3.2 Implementation in ANTARES

We restrict ourselves to the two-dimensional case for simplicity of presentation. The extension of the general formalism to the case of three spatial dimensions is straightforward. Rewriting the multispecies Navier–Stokes equations for the model problem leads to

$$\underbrace{\frac{d}{dt} \begin{pmatrix} \rho \\ \rho c \\ \rho \mathbf{u} \\ e \end{pmatrix}}_{\dot{y}(t)} = -\underbrace{\nabla \cdot \begin{pmatrix} \rho \mathbf{u} \\ \rho c \mathbf{u} \\ \rho \mathbf{u} \otimes \mathbf{u} \\ e \mathbf{u} \end{pmatrix}}_{F(y(t))} - \underbrace{\nabla \cdot \begin{pmatrix} 0 \\ 0 \\ P \\ P \mathbf{u} \end{pmatrix}}_{F(y(t))} + \underbrace{\nabla \cdot \begin{pmatrix} 0 \\ 0 \\ \sigma \\ \mathbf{u} \cdot \sigma \end{pmatrix}}_{F(y(t))} - \underbrace{\begin{pmatrix} 0 \\ 0 \\ \rho g \\ \rho g \mathbf{u} \end{pmatrix}}_{F(y(t))} + \underbrace{\nabla \cdot \begin{pmatrix} 0 \\ \rho \kappa_c \nabla c \\ 0 \\ \kappa \nabla T \end{pmatrix}}_{G(y(t))}. \quad (31)$$

The i^{th} implicit stage of the IMEX method is typically of the form

$$y_i = y^* + \Delta t \tilde{a}_{ii} G(y_i) \quad (32)$$

where y^* is known from previous stages. This translates to

$$\rho_i = \rho^*, \quad (33)$$

$$(\rho c)_i = (\rho c)^* + \Delta t \tilde{a}_{ij} \nabla \cdot (\rho_i \kappa_c \nabla c_i), \quad (34)$$

$$(\rho \mathbf{u})_i = (\rho \mathbf{u})^*, \quad (35)$$

$$e_i = e^* + \Delta t \tilde{a}_{ij} \nabla \cdot (\kappa \nabla T_i) \quad (36)$$

Rearranging (34) leads to

$$\frac{\rho^*}{\Delta t \tilde{a}_{ii}} c_i - \nabla \cdot (\rho^* \kappa_c \nabla c_i) = \frac{\rho^* c^*}{\Delta t \tilde{a}_{ii}} \quad (37)$$

Obviously, this is a general elliptic equation for c_i of the form

$$g(x, y) \varphi(x, y) - \nabla \cdot (h(x, y) \nabla \varphi(x, y)) = f(x, y). \quad (38)$$

Due to model assumptions, (36) can also be transformed to resemble a general elliptic equation. We start by recalling

$$e = e_{int} + e_{kin}, \quad (39)$$

$$= e_{int} + \frac{1}{2} \rho \mathbf{u}^2. \quad (40)$$

Bearing in mind equation (35) equation (36) reads

$$e_{int\ i} = e_{int}^* + \Delta t \tilde{a}_{ij} \nabla \cdot (\kappa \nabla T_i) \quad (41)$$

The equation of state for an ideal gas relates the temperature T and the internal energy e_{int} via

$$e_{int} = \frac{3}{2} \frac{T \rho R_{gas}}{m}, \quad (42)$$

if we assume the ratio of the specific heats at constant pressure and volume to equal 5/3. m denotes the mean molecular weight of the compound. Hence, we have at stage i

$$e_{int\ i} = \frac{3}{2} \frac{T_i \rho^* R_{gas}}{m_i} \quad (43)$$

and therefore we arrive at

$$\frac{3}{2} \frac{R_{gas} \rho^*}{m_i \Delta t \tilde{a}_{ii}} T_i - \nabla \cdot (\kappa_T \nabla T_i) = \frac{e_{int}^*}{\Delta t \tilde{a}_{ii}}. \quad (44)$$

Since m_i is evaluated using the mass fraction c_i , it is mandatory to solve equation (37) first.

Thus, the solution of an implicit stage translates to the solution of the generalized elliptic equations for the mass fraction c and the temperature T . In the ANTARES framework, finite elements are used for the discretization of (38). The resulting linear system is solved by the conjugate gradient method. For parallel computations, the Schur complement algorithm is applied. A detailed description is found in [8].

The above procedure applies without modification to the case of the fully compressible Navier–Stokes equation. However, for low Mach number flows a splitting approach is preferable where the terms containing pressure are treated separately in a post-processing step, i.e. after evaluation of all other terms for the computation of the velocity fields. The latter is obtained for the current step from an additional generalized Poisson equation for the pressure. This procedure is described in detail in [6]. Hence, the explicit stage is evaluated here using a fractional step method [21] as implemented in [11].

Timestep Control

Since sound waves are filtered according to the method of [21] as in [11], restrictions on the time step Δt are imposed by heat diffusion τ_T , diffusion of the second species τ_c , the viscosity τ_{visc} and the velocity of the fluid τ_{fluid} ,

$$\Delta t = \min\{\tau_c, \tau_T, \tau_{visc}, \tau_{fluid}\}, \quad (45)$$

where

$$\begin{aligned}\tau_c &= C_c \cdot 0.75 \cdot \frac{\min\{\Delta x^2, \Delta y^2\}}{\kappa_c} \\ \tau_T &= C_T \cdot 0.75 \cdot \frac{\min\{\Delta x^2, \Delta y^2\}}{\kappa_T} \\ \tau_{\text{visc}} &= C_{\text{visc}} \cdot 0.75 \cdot \frac{\min\{\Delta x^2, \Delta y^2\}}{\nu} \\ \tau_{\text{fluid}} &= C_{\text{fluid}} \frac{\min\{\Delta x, \Delta y\}}{\max(|\mathbf{u}|)}\end{aligned}$$

Usually, we define $C_c = C_T = C_{\text{fluid}} =: C_{\text{start}}$ and set $C_{\text{visc}} = 0.8$. The factor 0.75 in the calculation of τ_c , τ_T and τ_{visc} stems from the dissipative scheme (12) used to determine the parabolic parts of equations (26), (28) and (29), where our analysis has shown that a reduction of the time-step by this factor as compared to the standard three-point scheme (11) is necessary for stability and dissipativity of the discretization (see [18] and the results presented here in Sect. 2).

Since IMEX methods treat the diffusion term implicitly, the restrictions τ_c and τ_T can be neglected. This leads to the desired computational advantage. Generally, we use $\tau_{\text{diff}} = \min(\tau_c, \tau_T)$ as a starting value for our timestep, since initially $\tau_{\text{fluid}} \gg \tau_{\text{diff}}$.

The most effective criterion in regulating the timestep has been to watch out for two-point instabilities appearing in the conservative variables $(\rho, \rho c, \rho \bar{u}, Et)$. Due to the gravitational force operating vertically (note that in the ANTARES framework the positive x-axis is directed downwards), such instabilities are prone to appear in the horizontal direction.

To detect the occurrence of such oscillations we use the difference of the variables between two grid cells to determine the sign of the corresponding gradient. In case of the density ρ this reads

$$\begin{aligned}d_1 &= \rho_{i,j-1} - \rho_{i,j-2} \\ d_2 &= \rho_{i,j} - \rho_{i,j-1} \\ d_3 &= \rho_{i,j+1} - \rho_{i,j} \\ d_4 &= \rho_{i,j+2} - \rho_{i,j+1}\end{aligned}$$

where $1 < i < n_x$ and $1 < j < n_y$, assuming the grid consists of $n_x \times n_y$ grid points.

If the sign pattern of (d_1, d_2, d_3, d_4) corresponds to $(+, -, +, \pm)$, $(-, +, -, \pm)$, $(\pm, +, -, +,)$ or $(\pm, -, +, -)$, we have located a two-point instability.

The timestep control permits the occurrence of $n_y \cdot 0.1$ two-point instabilities for fixed i . If this limit is exceeded, the time step is repeated using a timestep of $\frac{2}{3}\tau$.

To permit the system to readjust after reducing the time step no modifications of τ are made for the next 15 timesteps regardless of the number of two-point instabilities. If the number of oscillations still exceeds the given limit after those 15 time steps, the timestep is again reduced.

If no or very few two-point oscillations are encountered for over 50 successive timesteps, the timestep is augmented by a factor of $\frac{5}{4}$.

Furthermore, the time step Δt must satisfy the following conditions:

- The time step restriction stemming from the viscous part provides an upper limit to Δt ,

$$\tau_{\text{visc}} \geq \Delta t, \quad (46)$$

whereas τ_{diff} serves as a lower limit to the timestep,

$$\tau_{\text{diff}} \leq \Delta t. \quad (47)$$

- In case of $\tau_{\text{diff}} \leq \tau_{\text{fluid}} \leq \tau_{\text{visc}}$, τ_{fluid} is an upper limit of Δt , whereas τ_{diff} provides a lower limit.

$$\tau_{\text{diff}} \leq \Delta t \leq \tau_{\text{fluid}} \quad (48)$$

- In case that the fluid velocity provides the most restrictive timestep limitation $\tau_{\text{fluid}} \leq \tau_{\text{diff}}$, we set $\Delta t = \tau_{\text{fluid}}$.

3.3 Simulations

In this section, we give the results of the simulations performed with the time integrators discussed in this report and compare their performance.

To define the test problem, according to [23] we specify a hydrostatic configuration with a temperature gradient which makes the fluid unstable against convection and which is counteracted by a gradient in mean molecular weight. The simulation of a single semiconvective layer requires the mean molecular weight to be linearly stratified. As time evolves, we expect convection to set in and mix the zone completely, although its development is inhibited by the stable mean molecular weight gradient.

Simulations have been performed on the Vienna Scientific Cluster, using 64 CPUs in parallel. The spatial resolution is 400×401 grid points. Simulation time is measured in units of *sound crossing times* (*scrt*). One *scrt* is defined as the time taking an acoustic wave to propagate from the bottom to the top of the simulation box. In our simulations, $1 \text{ scrt} = 5215.5 \text{ s}$ and the simulation time is 200 *scrt*.

Tables 14 and 15 sum up the performance achieved with the Runge–Kutta schemes presented. Since the capability of the method is best judged in that part of the simulation where the fluid velocity is too small to present a severe restriction on the timestep Δt , the CFL-numbers and timesteps in Table 14 have been measured in this regime. For Simulation 1, this corresponds to the first 80 *scrt* (see Figure 15).

We compare the largest possible time steps Δt_{\max} , the average time steps Δt_{mean} , the maximal and average CFL numbers resulting from the adaptive step selection, and the initial CFL numbers. The tests were performed for Prandtl numbers $\text{Pr} = 0.1$ and $\text{Pr} = 0.5$ distinguishing Simulations 1 and 2, respectively, a Lewis number $\text{Le} = 0.1$, $R_\rho = 1.1$, and a modified Rayleigh number $\text{Ra}^* = 160000$ related to the Rayleigh number through $\text{Ra}^* = \text{Ra} \cdot \text{Pr}$.

When we compare the performance of the second order schemes, it is obvious that the modification of γ in IMEX SSP2(2,2,2) has an astonishing effect on the stability of the scheme. The increased dissipativity of this method proves most effective in suppressing oscillations, permitting an average time step and CFL-number twice as high as the original method.

Table 14 shows that IMEX SSP2(3,3,2) permits a timestep and CFL number even three times higher than IMEX SSP2(2,2,2) with modified γ .

However, a comparison of the computation time given in Table 15 shows that the latter does not decrease by the same factor as the CFL-number, since as the time step Δt grows, the iterative solver used to solve the generalized elliptic problem requires more iterations to converge, resulting in an increase in computation time. A comparison of the computation times shows that, although IMEX SSP2(3,3,2) permits impressively large time steps, the need to solve three additional generalized Poisson problems related to the third stage takes its toll, whence the method’s performance is inferior to IMEX SSP2(2,2,2) with modified γ and in case of Simulation 2, is not even competitive to SSP(3,2), though it still performs superiorly to the O/S 2 scheme.

It is interesting to note that the initial CFL number for the first time step has a negligible influence on the actual CFL number reached in the diffusive part of the simulation. However, as soon as the fluid velocity implies a serious restriction on Δt , a higher initial CFL number leads to a significantly larger average time step and reduces the required computation time.

Note that optimization of the solver for the generalized Poisson equation (37) has the potential for a further significant decrease of computation time required.

A comparison of O/S 3 and IMEX SSP3(3,3,3) also shows that the larger timesteps lead to a significant gain in computational efficiency.

4 Comparisons

In this section, we give the analysis corresponding to Section 2 also for the classical Osher/Shu methods [29] (results taken from [18]) and the SDIRK methods of orders two and three with s stages [5]. Additionally, we give the results for the explicit scheme A from (19). We compare stability regions and dissipativity, where for the former we also formally use the definition based on (5). Figure 17 gives plots of the stability regions according to (5) for the forward Euler method, the Osher/Shu methods of second and third order [29], the SDIRK methods [4, 5] and the three IMEX schemes discussed in this report. The corresponding plots of the stability regions of the involved implicit methods \tilde{A} are given in Figure 18. Naturally, for the non-partitioned Runge–Kutta methods, the stability regions according to the IMEX definition and for the implicit schemes coincide. The stability regions in the IMEX setting could not be determined by MAPLE for SDIRK with $p = 3$, $s = 3, 4$. However, these can be replaced by the stability regions for the implicit methods. Note the different scaling that is applied for the plots of the stability regions of the third order SDIRK methods.

Table 16 sums up our complete analysis of the stability and dissipativity of all the methods in our focus. The first column gives the left boundaries of the stability regions of the implicit schemes \tilde{A} , the second column the left boundaries of the stability regions of the IMEX schemes, the third column gives the smallest positive point μ where the amplification factor for the respective time integrator in conjunction with spatial discretization by the three-point scheme becomes zero, the fourth column the point where the modulus of the amplification factor exceeds 1, columns 5 and 6 the corresponding data for the dissipative fourth order discretization [18] and the last column the error constant. An asterisk marks the zeros of g where the function undergoes a sign change. We note that the fully implicit methods have the best accuracy due to small error constants, and also excel in their stability and dissipativity properties, especially for a larger number of stages s . This must be contrasted with the higher

Method	Δt_{\max}	Δt_{mean}	CFL_{\max}	CFL_{mean}	$\text{CFL}_{\text{start}}$
Singlelayer $\text{Pr} = 0.1, \text{Le} = 0.1, R_\rho = 1.1, \text{Ra}^* = 160000$					
O/S 2	3.71 s	3.71 s	0.2	0.2	0.2
SSP(3,2)	9.31 s	9.31 s	0.5	0.5	0.5
IMEX SSP2(2,2,2)	22.21 s	11.56 s	1.20	0.62	0.2
IMEX SSP2(2,2,2), $\gamma=0.24$	36.35 s	19.44 s	1.96	1.05	0.2
IMEX SSP2(2,2,2), $\gamma=0.24$	36.35 s	19.43 s	1.96	1.05	0.3
IMEX SSP2(2,2,2), $\gamma=0.24$	37.02 s	19.45 s	2.00	1.05	0.4
IMEX SSP2(2,2,2), $\gamma=0.24$	35.53 s	19.47 s	1.91	1.05	0.5
IMEX SSP2(3,3,2)	74.52 s	57.37 s	4.02	3.09	0.4
IMEX SSP2(3,3,2)	93.15 s	57.16 s	5.02	3.08	0.5
O/S 3	3.71 s	3.71 s	0.2	0.2	0.2
IMEX SSP3(3,3,3)	15.14 s	10.14 s	0.82	0.55	0.2
Singlelayer $\text{Pr} = 0.5, \text{Le} = 0.1, R_\rho = 1.1, \text{Ra}^* = 160000$					
O/S 2	3.72 s	3.72 s	0.2	0.2	0.2
SSP(3,2)	9.31 s	9.31 s	0.5	0.5	0.5
IMEX SSP2(2,2,2)	23.14 s	13.84 s	1.24	0.74	0.2
IMEX SSP2(2,2,2), $\gamma=0.24$	23.13 s	15.72 s	1.24	0.85	0.2
IMEX SSP2(2,2,2), $\gamma=0.24$	22.76 s	15.79 s	1.22	0.85	0.3
IMEX SSP2(2,2,2), $\gamma=0.24$	20.32 s	15.81 s	1.09	0.85	0.4
IMEX SSP2(2,2,2), $\gamma=0.24$	22.81 s	15.85 s	1.23	0.84	0.5
IMEX SSP2(3,3,2)	40.65 s	33.54 s	2.19	1.80	0.4
IMEX SSP2(3,3,2)	40.65 s	33.87 s	2.19	1.82	0.5
O/S 3	3.72 s	3.72 s	0.2	0.2	0.2
IMEX SSP3(3,3,3)	15.05 s	9.70 s	0.81	0.52	0.2

Table 14: Numerical comparisons: Timesteps and CFL-numbers

Method	$\text{CFL}_{\text{start}}$	computation time	number of time steps
Singlelayer $\text{Pr} = 0.1, \text{Le} = 0.1, R_\rho = 1.1, \text{Ra}^* = 160000$			
O/S 2	0.2	7:17:32	287 788
SSP(3,2)	0.2	4:40:24	115 691
IMEX SSP2(2,2,2)	0.2	6:04:54	92 492
IMEX SSP2(2,2,2), $\gamma=0.24$	0.2	4:41:15	63 586
IMEX SSP2(2,2,2), $\gamma=0.24$	0.3	4:18:59	56 741
IMEX SSP2(2,2,2), $\gamma=0.24$	0.4	4:10:26	54 015
IMEX SSP2(2,2,2), $\gamma=0.24$	0.5	4:12:38	53 941
IMEX SSP2(3,3,2)	0.4	4:31:12	27 673
IMEX SSP2(3,3,2)	0.5	4:27:46	24 266
O/S 3	0.2	10:55:40	288 607
IMEX SSP3(3,3,3)	0.2	8:19:39	104 789
Singlelayer $\text{Pr} = 0.5, \text{Le} = 0.1, R_\rho = 1.1, \text{Ra}^* = 160000$			
O/S 2	0.2	7:01:44	286 011
EXP 332	0.5	4:34:35	114 409
IMEX SSP2(2,2,2)	0.2	5:10:01	75 384
IMEX SSP2(2,2,2), $\gamma=0.24$	0.2	4:45:36	66 997
IMEX SSP2(2,2,2), $\gamma=0.24$	0.3	4:42:00	68 591
IMEX SSP2(2,2,2), $\gamma=0.24$	0.4	4:42:24	66 847
IMEX SSP2(2,2,2), $\gamma=0.24$	0.5	4:50:52	66 828
IMEX SSP2(3,3,2)	0.4	4:43:27	31 225
IMEX SSP2(3,3,2)	0.5	4:43:26	31 112
O/S 3	0.2	10:45:43	286 006
IMEX SSP3(3,3,3)	0.2	7:27:01	108 852

Table 15: Numerical comparisons: computation times and overall number of timesteps

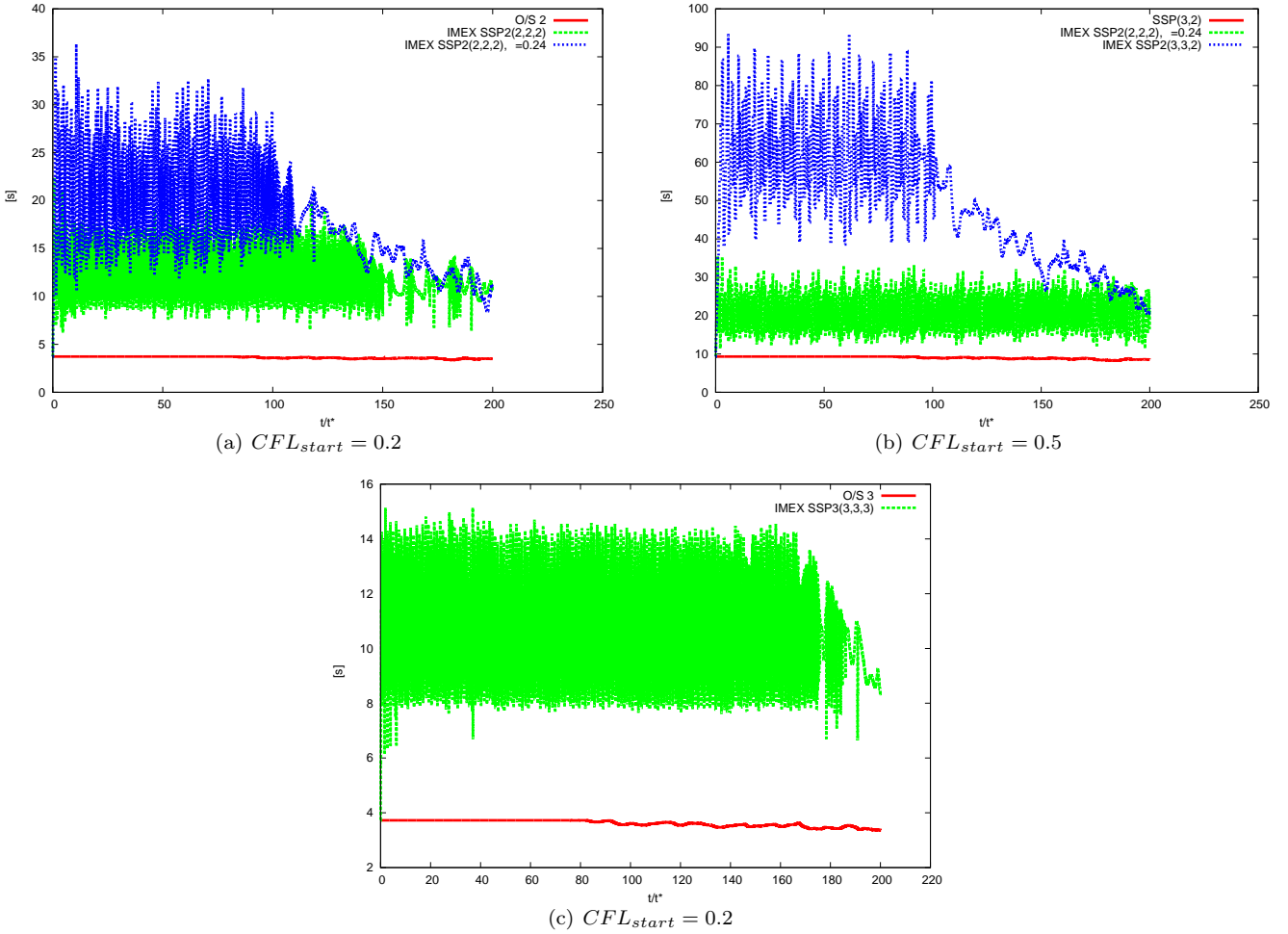


Figure 15: Timestep evolution in Simulation 1. Pictures (a) and (b) compare the timestep of the second order schemes whereas (c) shows the evolution of Δt using O/S 3 and IMEX SSP3(3,3,3). Integration interval $t/t^* \in [0, 200]$ with $t^* = 5215.5 \text{ s} = 1 \text{ scrt}$.

computational effort and step size limitations resulting from the need to solve the model equations by fixed point iteration [18]. Finally, we stress that in comparison with the other explicit methods, the explicit sub-scheme A from (19) seems to have the most favorable properties. As compared to the IMEX and the fully implicit SDIRK methods, there is a trade-off between accuracy, stability, dissipativity and the numerical effort for the realization of the schemes.

Acknowledgements

The authors thank I. Higuera for providing the analysis of method (13) for variable values of γ [12].

References

- [1] U. Ascher, S. Ruuth, and R. Spiteri. Implicit–explicit Runge–Kutta methods for time-dependent partial differential equations. *Appl. Numer. Math.*, 25:151–167, 1997.
- [2] U. Ascher, S. Ruuth, and T.R. Wetton. Implicit–explicit methods for time-dependent partial differential equations. *SIAM J. Numer. Anal.*, 32(3):797–823, 1995.
- [3] R. Donat and A. Marquina. Capturing shock reflections: An improved flux formula. *J. Comput. Phys.*, 125:42–58, 1996.
- [4] L. Ferracina. *Monotonicity and Boundedness in General Runge–Kutta Methods*. Ph.D. Thesis, Leiden University, Leiden, The Netherlands, 2005.
- [5] L. Ferracina and M. Spijker. Strong stability of singly-diagonally-implicit Runge–Kutta methods. *Appl. Numer. Math.*, 58:1675–1686, 2008.

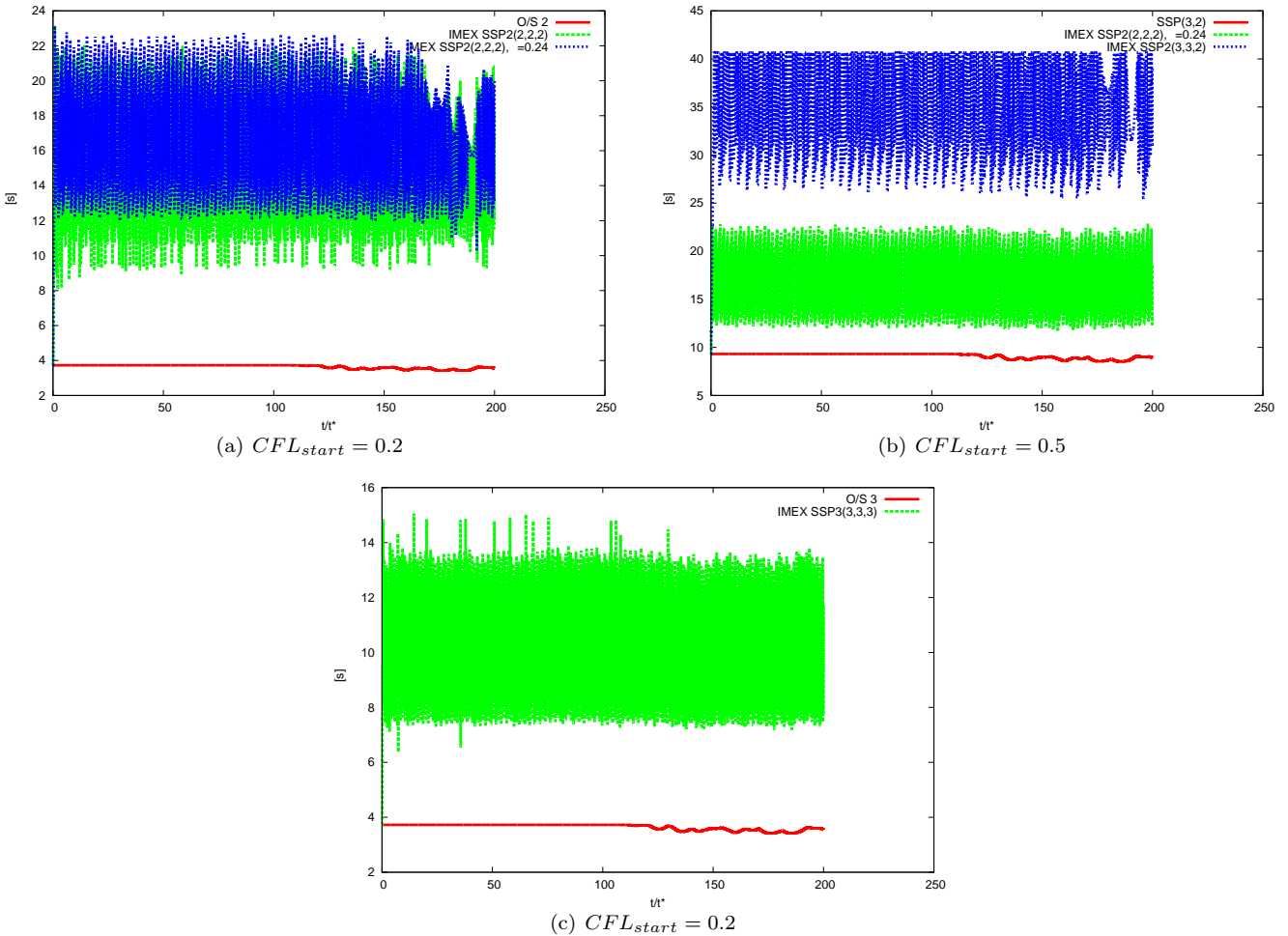


Figure 16: Timestep evolution in Simulation 2. Pictures (a) and (b) compare the timestep of the second order schemes whereas (c) shows the evolution of Δt using O/S 3 and IMEX SSP3(3,3,3). Integration interval $t/t^* \in [0, 200]$ with $t^* = 5215.5 \text{ s} = 1 \text{ scrt}$.

- [6] J.H. Ferziger and Perić. *Computational Methods for Fluid Dynamics*. Springer Verlag, Berlin–Heidelberg–New York, 2002.
- [7] S. Gottlieb and L.-A. Gottlieb. Strong stability preserving properties of Runge–Kutta time discretization methods for linear constant coefficient operators. *J. Sci. Comput.*, 18:83–109, 2003.
- [8] H. Grimm-Strele. Numerical solution of the generalised Poisson equation on parallel computers. Master’s thesis, University of Vienna, March 2010. Available from <http://othes.univie.ac.at/9200/>.
- [9] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I*. Springer-Verlag, Berlin–Heidelberg–New York, 1987.
- [10] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II*. Springer-Verlag, Berlin–Heidelberg–New York, 1991.
- [11] N. Happenhofer. Simulation of low mach number fluids. Master’s thesis, University of Vienna, June 2010. Available from <http://othes.univie.ac.at/10551/>.
- [12] I. Higueras. Radius and regions for arbitrary γ . Private communication.
- [13] I. Higueras. Representations of Runge–Kutta methods and strong stability preserving methods. *SIAM J. Numer. Anal.*, 43:924–948, 2005.
- [14] I. Higueras. Strong stability for additive Runge–Kutta methods. *SIAM J. Numer. Anal.*, 44:1735–1758, 2006.
- [15] I. Higueras. Characterizing strong stability preserving additive Runge–Kutta methods. *J. Sci. Comput.*, 39:115–128, 2009.
- [16] C.A. Kennedy and M.H. Carpenter. Additive Runge–Kutta schemes for convection–diffusion–reaction equations. *Appl. Numer. Math.*, 44:139–181, 2003.

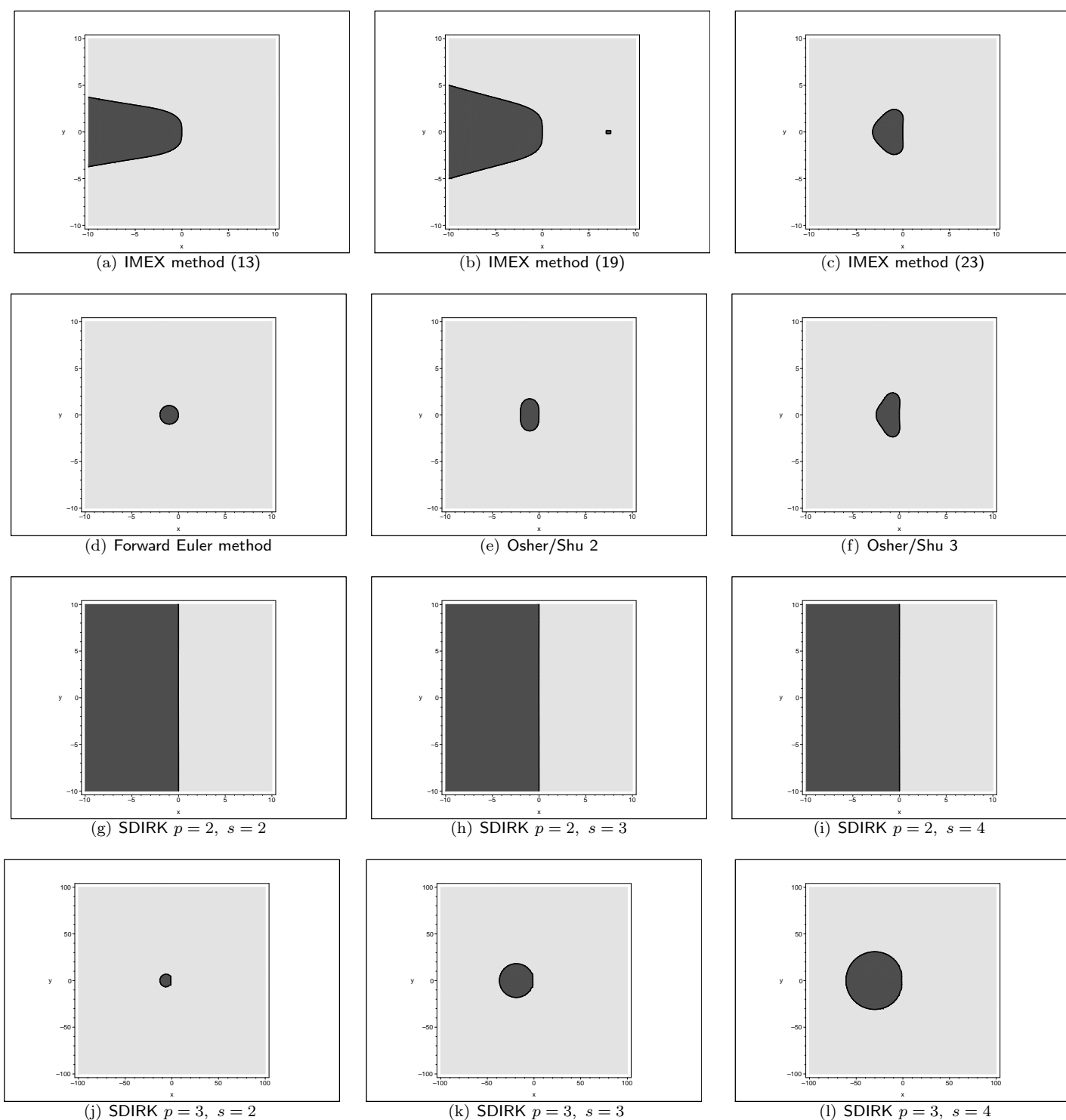


Figure 17: Comparison of stability regions for IMEX methods.

- [17] R. Kippenhahn and A. Weigert. *Stellar Structure and Evolution*. Springer-Verlag, 3rd printing of 1st edition, 1994.
- [18] O. Koch, F. Kupka, B. Löw-Baselli, A. Mayrhofer, and F. Zaussinger. SDIRK methods for the ANTARES code. ASC Report 32/2010, Inst. for Anal. and Sci. Comput., Vienna Univ. of Technology, 2010. Available at <http://www.asc.tuwien.ac.at/preprint/2010/asc32x2010.pdf>.
- [19] T. Koto. IMEX Runge–Kutta schemes for reaction–diffusion equations. *J. Comput. Appl. Math.*, 215:182–195, 208.
- [20] J.F.B.M. Kraaijevanger. Contractivity of Runge–Kutta methods. *BIT*, 31:482–528, 1991.

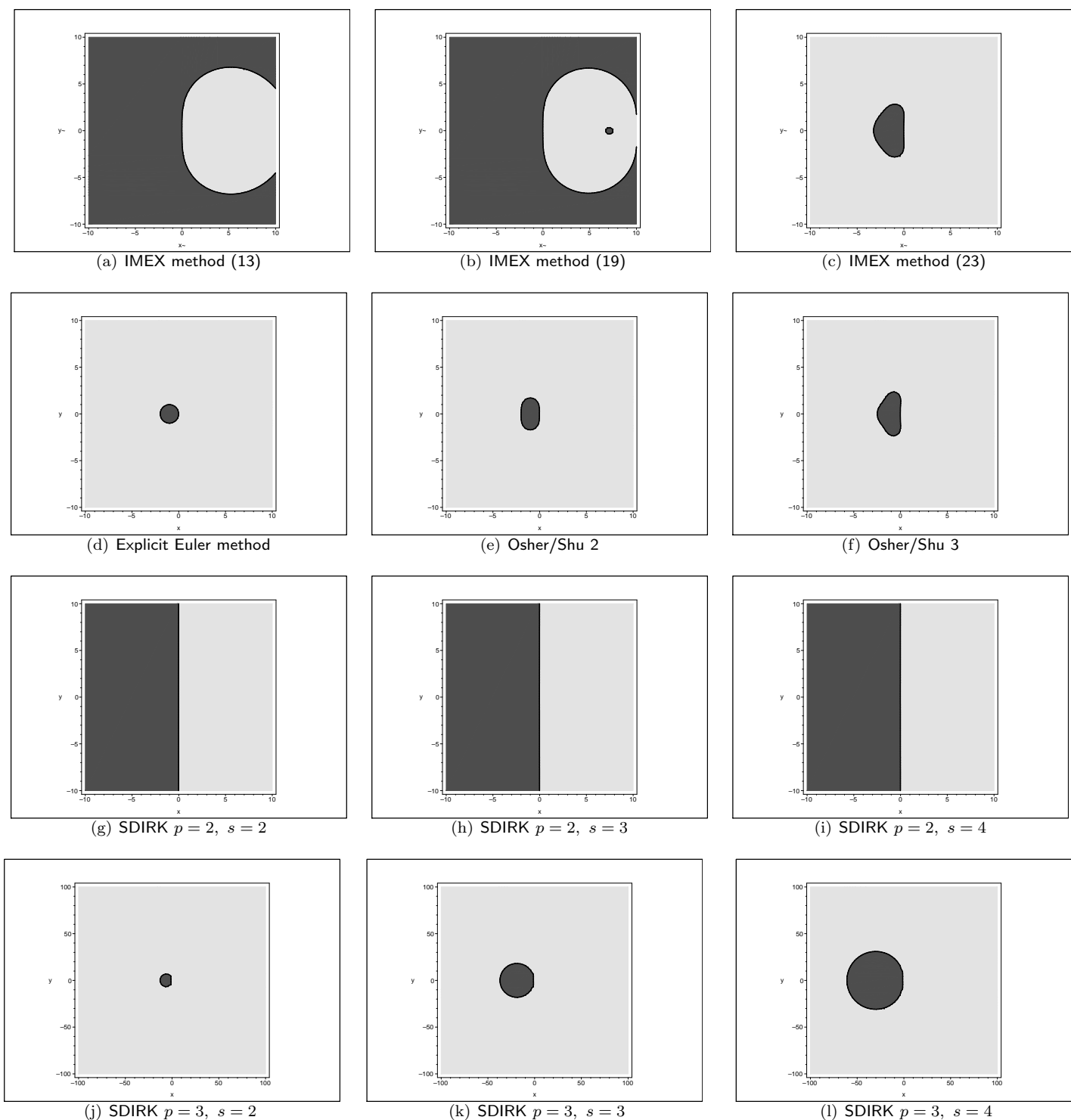


Figure 18: Comparison of stability regions for implicit schemes \tilde{A} .

- [21] N. Kwatra, J. Su, J.T. Grétarsson, and R. Fedkiw. A method for avoiding the acoustic time step restriction in compressible flow. *J. Comput. Phys.*, 228:4146–4161, June 2009.
- [22] X. Liu and S. Osher. Convex ENO high order multi-dimensional schemes without field by field decomposition or staggered grids. *J. Comput. Phys.*, 142:304–330, 1998.
- [23] H.J. Muthsam, W. Göb, F. Kupka, and W. Liebich. Interacting convection zones. *New Astronomy*, 4:405–417, 1999.
- [24] H.J. Muthsam, F. Kupka, B. Löw-Baselli, C. Obertscheider, M. Langer, and P. Lenz. ANTARES — A Numerical Tool for Astrophysical RESearch with applications to solar granulation. *New Astronomy*, 15:460–475, 2010.

Method	Stability \tilde{A}	Stability IMEX	$g_{3\text{pt}} = 0$	$ g_{3\text{pt}} = 1$	$g_{4\text{th}} = 0$	$ g_{4\text{th}} = 1$	C
IMEX (13)	$-\infty$	$-\infty$	0.603*	—	0.452*	—	5.17
IMEX (13), $\gamma = 0.24$	-50	-50	—	12.50	—	9.375	2.79
IMEX (19)	$-\infty$	$-\infty$	0.606*	—	0.455*	—	8.05
IMEX (23)	-3.248	-3.248	0.465*	0.812	0.348*	0.609	11.6
Forward Euler	-2	-2	0.25*	0.5	0.187*	0.375	12.6
Osher/Shu 2	-2	-2	—	0.5	—	0.375	16.2
Osher/Shu 3	-2.512	-2.512	0.399*	0.628	0.299*	0.471	22.8
SSP(3,2)	-4.519	-4.519	0.896*	1.129	0.672*	0.847	6.40
SDIRK $p = 2, s = 2$	$-\infty$	$-\infty$	1	—	0.75	—	1.64
SDIRK $p = 2, s = 3$	$-\infty$	$-\infty$	1.5*	—	1.125*	—	0.72
SDIRK $p = 2, s = 4$	$-\infty$	$-\infty$	2	—	1.5	—	0.40
SDIRK $p = 3, s = 2$	-12.93	-12.93	—	3.232	—	2.424	1.43
SDIRK $p = 3, s = 3$	-37.10	-37.10	2.328*	9.237	1.746*	6.955	0.56
SDIRK $p = 3, s = 4$	-60.98	-60.98	—	15.25	—	11.44	0.30

Table 16: Summary of the analysis of TVD integrators.

- [25] L. Pareschi and G. Russo. Implicit–explicit Runge–Kutta schemes and application to hyperbolic systems with relaxation. *J. Sci. Comput.*, 25:129–155, 2005.
- [26] S.J. Ruuth and R.J. Spiteri. High-order strong-stability-preserving runge-kutta methods with downwind-biased spatial discretizations. *SIAM J. Numer. Anal.*, 42:974–996, 2004.
- [27] C.-W. Shu. Total-variation-diminishing time discretizations. *SIAM J. Sci. Statist. Comput.*, 9:1073–1084, 1988.
- [28] C.-W. Shu. Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. Technical Report ICASE 97-65, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA, 1997.
- [29] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.*, 77:439–471, 1988.
- [30] R.J. Spiteri and S.J. Ruuth. A new class of optimal high-order strong-stability-preserving time discretization methods. *SIAM J. Numer. Anal.*, 40:469–491, 2002.
- [31] J.C. Strikwerda. *Finite Difference Schemes and Partial Differential Equations*. SIAM, Philadelphia, PA, 2nd edition, 2004.
- [32] A. Weiss, W. Hillebrandt, H.C. Thomas, and H. Ritter. *Cox and Giuli’s Principles of Stellar Structure*. Advances in Astronomy and Astrophysics. Cambridge Scientific Publishers Ltd., Cambridge, U.K., 2nd edition, 2004.
- [33] F. Zaussinger and H.C. Spruit. Semiconvection. Submitted to *Astron. Astrophys.*, 2011.