AN ANALYSIS OF THE PROTHERO–ROBINSON EXAMPLE FOR CONSTRUCTING NEW DIRK AND ROW METHODS
This document was created April 2012 using \LaTeX\ 2\epsilon.

Institute of Scientific Computing
Technische Universität Braunschweig
Hans-Sommer-Straße 65
D-38106 Braunschweig, Germany

url: www.wire.tu-bs.de
mail: wire@tu-bs.de

Copyright © by Joachim Rang

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted in connection with reviews or scholarly analysis. Permission for use must always be obtained from the copyright holder.

Alle Rechte vorbehalten, auch das des auszugsweisen Nachdrucks, der auszugsweisen oder vollständigen Wiedergabe (Photographie, Mikroskopie), der Speicherung in Datenverarbeitungsanlagen und das der Übersetzung.
An analysis of the Prothero–Robinson example for constructing new DIRK and ROW methods

Joachim Rang
Institute of Scientific Computing, TU Braunschweig
j.rang@tu-bs.de

Abstract

In this note the order reduction phenomenon of diagonally implicit Runge-Kutta methods (DIRK–methods) and Rosenbrock–Wanner methods (ROW–methods) applied on the Prothero-Robinson example is analysed. New order conditions to avoid order reduction are derived and a new second order DIRK and ROW–method is created. The new schemes are applied on the Prothero–Robinson example and on the semi-discretised incompressible Navier–Stokes equations. Numerical examples show that the new methods converge with second order for velocity and pressure.

Keywords: ODEs, order reduction, DIRK methods, ROW methods
1 Introduction

One possibility to solve stiff ODEs like the example of Prothero and Robinson [15] or differential algebraic equations are Runge-Kutta methods [4, 23]. Explicit methods may not be a good choice since for getting a stable numerical solution a stepsize restriction should be accepted, i.e. the problem should be solved with very small timesteps. Therefore it might be better to use implicit methods, as for example Runge-Kutta methods, or linear-implicit methods such as Rosenbrock–Wanner methods. But in these cases the convergence may not be achieved [4, 23], i.e. the so-called order reduction phenomenon can be observed. In [4] convergence results for implicit Runge–Kutta methods can be found where the so-called stage order plays an important role. Ostermann and Roche prove in [13] that implicit Runge–Kutta methods may have a fractional order of convergence for general linear ODEs. Similar results are presented for Rosenbrock–Wanner methods in [14]. As for diagonally implicit Runge–Kutta methods with non-zero diagonal entries Rosenbrock–Wanner methods can have only stage order 1. That is the reason why Ostermann and Roche derive further order conditions for Rosenbrock–Wanner to reduce order reduction. For example in [11] and [18] Rosenbrock-Wanner methods are derived which satisfy the order conditions from Ostermann and Roche [14] and which have almost no order reduction if they are applied on stiff ODEs as the Prothero–Robinson example or the semi-discretised Navier-Stokes equations [18, 19, 8, 9]. In [21] a different approach can be found for reducing the order reduction. A Rosenbrock–Wanner method satisfying the order conditions derived by Scholz [21] is the RODASP method from Steinebach [22].

Fully implicit Runge–Kutta methods may be ineffective for solving high dimensional ODEs (as for example the incompressible Navier-Stokes equations) since they need a high computational effort. In this case it might be better to use diagonally implicit Runge–Kutta methods or Rosenbrock–Wanner methods (see for example [9]). But in this case the stage order is limited by 2 for the diagonally implicit Runge–Kutta methods (in the case of a singular coefficient matrix $A$). One question which this note tries to answer is the following: Can we construct a DIRK method with stage order 1 which converges with order 2 in the stiff case if applied for example on the ODE of Prothero and Robinson [15]?

The following considerations are motivated by the following observation. In [16] an embedded method for the fractional step-$\theta$ scheme (a special diagonally implicit Runge–Kutta method) is introduced and it is shown that this method has the convergence order 1. Solving the stiff Prothero-Robinson example it can be observed that the method converges with order 2 although
the known theoretical results suggest a convergence order of 1.

In this note we consider diagonally implicit and Rosenbrock–Wanner methods and apply them on the Prothero–Robinson example. In Section 3 we are considering the local error of these classes of methods in the non-stiff and in the stiff case. We will see that we get further order conditions which are needed to decrease the order reduction. A second order diagonally implicit Runge–Kutta method and second order Rosenbrock–Wanner method are created in Section 4, and finally we present some numerical results and apply our new methods on the Prothero–Robinson example and the incompressible Navier-Stokes equations.

2 Time discretisation

It is well-known that one-step methods have order reduction if they are applied on the Prothero–Robinson example, i.e. problem (4) and if $\lambda \ll 0$. Therefore we determine the local error of the ROW–methods and of the DIRK–methods with a regular coefficient matrix $A$.

2.1 Rosenbrock–Wanner methods

Application to ODEs. First we consider an ODE of the form

$$\dot{u} = F(t, u), \quad u(0) = u_0. \quad (1)$$

A Rosenbrock–Wanner–method (ROW–method) with $s$ internal stages is given by

$$MK_i = F\left(t_m + \alpha_i \tau_m, \tilde{U}_i\right) + \tau_m J \sum_{j=1}^{i} \gamma_{ij} k_j + \tau_m \gamma_i \dot{F}(t_m, u_m), \quad (2)$$

$$\tilde{U}_i = u_m + \tau_m \sum_{j=1}^{i-1} a_{ij} k_j, \quad i = 1, \ldots, s,$$

$$u_{m+1} = u_m + \tau_m \sum_{i=1}^{s} b_i k_i, \quad (3)$$

where $J := \partial_u F(t_m, u_m), \alpha_{ij}, \gamma_{ij}, b_i$ are the parameters of the method,

$$\alpha_i := \sum_{j=1}^{i-1} \alpha_{ij}, \quad \gamma_i := \sum_{j=1}^{i-1} \gamma_{ij}, \quad \gamma := \gamma_{ii} > 0, \quad i = 1, \ldots, s.$$
If the parameters $\alpha_{ij}$, $\gamma_{ij}$, and $b_i$ are chosen appropriately, a sufficient consistency order can be obtained. Additional consistency conditions arise if $J$ is only an approximation to $\partial_u F(t_m, \mathbf{u}_m)$, or if $J$ is an arbitrary matrix. This class of methods is called W–methods, [23]. If a ROW–method is applied to a semidiscretised partial differential equation, further order condition should be satisfied to avoid order reduction, see [12].

The ROW–method (2)–(3) requires the successive solution of $s$ linear systems of equations with the same matrix $I - \gamma \tau J$. The right hand side of the $i$–th linear system of equations depends on the solutions of the first to the $(i-1)$–st system. Thus, a main difference of ROW–methods to implicit methods is that it is not necessary to solve a nonlinear system of equations in each discrete time but only a fixed number of linear systems of equations.

**Application to the example of Prothero–Robinson.** In this section we apply the ROW–method (2)–(3) on the Prothero–Robinson problem, i. e. on

\[
\dot{u} = \lambda (u - \varphi(t)) + \dot{\varphi}(t), \quad u(0) = \varphi(0), \lambda < 0. \tag{4}
\]

The exact solution is given by $u(t) = \varphi(t)$. Inserting (4) into (2) yields

\[
k_i = \lambda \left( u_m + \tau \sum_{j=1}^{i} \beta_{ij} k_j - \varphi(t_m + \alpha_i \tau) \right) + \dot{\varphi}(t_m + \alpha_i \tau) + \tau \gamma_i (\ddot{\varphi}(t_m) - \lambda \dot{\varphi}(t_m)),
\]

where $\beta_{ij} := \alpha_{ij} + \beta_{ij}$. To abbreviate we set

\[
\varphi^{(k)}_i := \varphi^{(k)}(t_m + \alpha_i \tau), \quad i = 1, \ldots, s, k \geq 0,
\]

\[
\varphi^{(k)}_m := \varphi^{(k)}(t_m), \quad k \geq 0,
\]

\[
\mathbf{\Phi}^{(k)} := (\varphi^{(k)}_1, \ldots, \varphi^{(k)}_s)^\top, \quad \mathbf{k} := (k_1, \ldots, k_s)^\top, \quad \mathbf{e} := (1, \ldots, 1)^\top,
\]

\[
\alpha := (\alpha_1, \ldots, \alpha_s)^\top, \quad \gamma := (\gamma_1, \ldots, \gamma_s)^\top.
\]

It follows

\[
k_i = \lambda \left( u_m + \tau \sum_{j=1}^{i} \beta_{ij} k_j - \varphi_i \right) + \dot{\varphi}_i + \tau \gamma_i (-\lambda \dot{\varphi}_m + \ddot{\varphi}_m).
\]

Using the vector notation introduced above we obtain

\[
\mathbf{k} = \lambda (u_m \mathbf{e} + \tau B \mathbf{k} - \mathbf{\Phi}) + \dot{\mathbf{\Phi}} + \tau \gamma (\ddot{\varphi}_m - \lambda \dot{\varphi}_m)
\]

and

\[
\mathbf{k} = (I - zB)^{-1} (\lambda u_m \mathbf{e} - \lambda \mathbf{\Phi} + \dot{\mathbf{\Phi}} + \tau \gamma (\ddot{\varphi}_m - \lambda \dot{\varphi}_m)), \tag{5}
\]
where \( z := \lambda \tau \). Inserting (5) into (3) yields

\[
\begin{align*}
    u_{m+1} &= u_m + \tau b^\top (I - zB)^{-1} \left[ \lambda (u_me - \Phi) + \dot{\Phi} + \tau \gamma (\ddot{\varphi}_m - \lambda \dot{\varphi}_m) \right] \\
    &= u_m + z b^\top (I - zB)^{-1} [u_me - \Phi - \tau \gamma \dot{\varphi}_m] \\
    &\quad + \tau b^\top (I - zB)^{-1} [\dot{\Phi} + \tau \gamma \ddot{\varphi}_m].
\end{align*}
\]  

(6)

**Adaptive time step control.** ROW–methods have the advantage that they allow an easy implementation of an adaptive time steplength control. Consider a ROW–method of order \( p \geq 2 \). An adaptive time step control employs a second ROW–method which has the coefficients \( a_{ij}, \hat{b}_i, \) and \( c_i \), \( i, j = 1, \ldots, s \), and order \( p - 1 \). The solution of the second method at \( t_{m+1} \) is given by

\[
\hat{u}_{m+1} = u_m + \sum_{i=1}^{s} \hat{b}_i k_i.
\]

Now, the next time step \( \tau_{m+1} \) is proposed to be

\[
\tau_{m+1} = \rho \frac{\tau_{m}^2}{\tau_{m-1}} \left( \frac{TOL \cdot r_m}{r_{m+1}^2} \right)^{1/p},
\]

(7)

where \( \rho \in (0, 1] \) is a safety factor, \( TOL > 0 \) is a given tolerance and

\[
r_{m+1} := \| u_{m+1} - \hat{u}_{m+1} \|.
\]

(8)

This step size selection rule is called PI–controller and going back to Gustafsson et. al. [3]. For details on the numerical error and the implementation of automatic steplength control we refer to [4, 10].

### 2.2 DIRK–methods

**Application to ODEs.** As in the case of DIRK schemes we start our considerations with the ODE (1). A Runge–Kutta method (RK method) with \( s \) internal stages [4, 23] is a one–step–method for solving (1) of the form

\[
M k_i = F \left( t_m + \alpha_i \tau_m, U_i \right), \quad U_i = u_m + \tau_m \sum_{j=1}^{s} a_{ij} k_j, \quad i = 1, \ldots, s,
\]

(9)

\[
u_{m+1} = u_m + \tau_m \sum_{i=1}^{s} b_i k_i.
\]

(10)

The coefficients \( a_{ij}, b_i \) and \( c_i \) should be chosen in such a way that some order conditions are satisfied to obtain a sufficient consistency order.
In this paper the coefficients of the RK–method (9)–(10) satisfy \( a_{ij} = 0 \) for \( i < j, i, j \in \{1, \ldots, s\} \) and \( a_{ii} \neq 0 \) for \( i \in \{1, \ldots, s\} \). RK–methods satisfying these conditions are called diagonal–implicit RK–methods (DIRK–methods). These methods are discussed in several papers and books, e.g. in [23, 4].

**Application to the example of Prothero–Robinson.** Next we apply the DIRK–method (9)–(10) on the ODE (4). Inserting yields
\[
k_i = \lambda \left( u_m + \tau \sum_{j=1}^{i} \alpha_{ij} k_j - \varphi(t_m + \alpha_i \tau) \right) + \dot{\varphi}(t_m + \alpha_i \tau)
\]
and with the above defined abbreviations it follows
\[
k_i = \lambda \left( u_m + \tau \sum_{j=1}^{i} a_{ij} k_j - \varphi_i \right) + \dot{\varphi}_i.
\]
With the vector notation introduced above and the setting \( A = (a_{ij})_{i,j=1}^{s} \) it follows
\[
k = \lambda(u_m e + \tau A k - \Phi) + \dot{\Phi}
\]
and
\[
k = (I - zA)^{-1}(\lambda u_m e - \lambda \Phi + \dot{\Phi}). \tag{11}
\]
Inserting into (11) into (10) yields
\[
u_{m+1} = u_m + \tau b^\top (I - zA)^{-1}[(\lambda u_m e - \Phi) + \dot{\Phi}]
= u_m + z b^\top (I - zA)^{-1}[u_m e - \Phi] + \tau b^\top (I - zA)^{-1} \dot{\Phi}. \tag{12}
\]
Again, as in the last section about ROW–methods, an automatic step length control can be implemented with the help of an embedded method which exist in common DIRK–methods [9].

### 3 New order conditions

If we set \( B = A \) and \( \gamma = 0 \) in (6) then we get (12). Therefore, we analyse ROW–methods in the following. Next we compute the local error of the ROW–method (2)–(3) if it is applied on the Prothero–Robinson example. We have
\[
\delta_{\tau}(t_{m+1}) = u_{m+1} - \varphi(t_{m+1})
= \varphi_m - \varphi_{m+1} + z b^\top (I - zB)^{-1}[(\varphi_m e - \Phi - \tau \gamma \dot{\varphi}_m]
+ \tau b^\top (I - zB)^{-1}[\dot{\Phi} + \tau \gamma \ddot{\varphi}_m]. \tag{13}
\]
In the non-stiff case, we have $z \to 0$ for $\tau \to 0$, but in the stiff case $z$ tends to infinity if $\tau \to 0$.

### 3.1 The local error in the non-stiff case

In the non-stiff case, we can expand the term $(I - zB)^{-1}$ for small $\tau$ as follows:

$$(I - \lambda \tau B)^{-1} = I + \lambda B + (\lambda B)^2 + (\lambda B)^3 + \ldots.$$  

For the expansion of numerical solution $u_{m+1}$, we need the derivative of $fg$, where $f, g : \mathbb{R} \to \mathbb{R}$ should be sufficiently smooth. Then

$$\frac{d^k(fg)}{dx^k}(x) = \sum_{i=0}^{k} \binom{k}{i} \frac{d^i f(x)}{dx^i} \frac{d^{k-i} g(x)}{dx^{k-i}}, \quad k = 1, 2, \ldots.$$  

The Taylor expansion of $\delta_\tau(t_{m+1})$ reads as

$$\delta_\tau(t_{m+1}) = z \sum_{k=0}^{p-1} b^\top \sum_{l=0}^{k} \binom{k}{l} l!(\lambda B)^l \left[ \varphi_m e^{\delta_{k-l,0}} - \alpha^{k-l} \varphi_m^{(k-l)} - \gamma \lambda \varphi_m^{\delta_{k-l,1}} \right] \frac{\tau^k}{k!}$$

$$+ \tau \sum_{k=0}^{p-1} b^\top \sum_{l=0}^{k} \binom{k}{l} l!(\lambda B)^l \left[ \alpha^{k-l} \varphi_m^{(k-l+1)} + \gamma \varphi_m^{\delta_{k-l,1}} \right] \frac{\tau^k}{k!}$$

$$+ \varphi_m - \sum_{k=0}^{p} \varphi_m^{(k)} \frac{\tau^k}{k!} + O(\tau^{p+1}).$$

For $k = l$ the term in the first row drops out and it follows with the definition of the binomial coefficient and a splitting of the sum in the second row

$$\delta_\tau(t_{m+1}) = -z \sum_{k=1}^{p-1} b^\top \sum_{l=0}^{k-1} (\lambda B)^l \left[ \alpha^{k-l} + \gamma \delta_{k-l,1} \right] \varphi_m^{(k-l)} \frac{\tau^k}{(k-l)!}$$

$$+ \sum_{k=1}^{p-1} b^\top \sum_{l=1}^{k} (\lambda B)^l \left[ \alpha^{k-l} + \gamma \delta_{k-l,1} \right] \varphi_m^{(k-l+1)} \frac{\tau^{k+1}}{(k-l)!}$$

$$+ \sum_{k=0}^{p-1} b^\top \left[ \alpha^k + \gamma \delta_{k,1} \right] \frac{\tau^{k+1}}{k!} \varphi_m^{(k+1)} - \sum_{k=1}^{p} \varphi_m^{(k)} \frac{\tau^k}{k!} + O(\tau^{p+1}).$$
Next we sum up in the second row from $l = 0$ to $k - 1$ and get

$$
\delta_\tau(t_{m+1}) = -z \sum_{k=1}^{p-1} \sum_{l=1}^{k-1} (\lambda B)^l \left[ \alpha^{k-l} + \gamma \delta_{k-l,1} \right] \varphi_m^{(k-l)} \frac{\tau^k}{(k-l)!} \\
+ \sum_{k=1}^{p-1} \sum_{l=0}^{k-1} (\lambda B)^{l+1} \left[ \alpha^{k-l-1} + \gamma \delta_{k-l-1,1} \right] \varphi_m^{(k-l)} \frac{\tau^{k+1}}{(k-l-1)!} \\
+ \sum_{k=0}^{p-1} \left[ \alpha^k + \gamma \delta_{k,1} \right] \frac{\tau^{k+1}}{k!} \varphi_m^{(k+1)} - \sum_{k=0}^{p-1} \varphi_m^{(k+1)} \frac{\tau^{k+1}}{(k+1)!} + O(\tau^{p+1}) \\
= z \sum_{k=1}^{p-1} \sum_{l=0}^{k-2} (\lambda B)^l \left\{ B \left[ \alpha^{k-l-1} + \gamma \delta_{k-l-1,1} \right] - \frac{1}{k-l} \alpha^{k-l} \right\} \\
\cdot \varphi_m^{(k-l)} \frac{\tau^k}{(k-l-1)!} \\
+ \sum_{k=0}^{p-1} \left\{ \alpha^k + \gamma \delta_{k,1} \right\} \varphi_m^{(k)} \frac{\tau^k}{k!} + O(\tau^{p+1}),
$$

where we use the fact that

$$
B \left[ \alpha^{k-l-1} + \gamma \delta_{k-l-1,1} \right] - \frac{1}{k-l} \left[ \alpha^{k-l} + \gamma \delta_{k-l,1} \right] = 0,
$$

if $l = k - 1$. For Rosenbrock methods we get the order conditions

$$
b^\top (\alpha^k + \gamma \delta_{k,1}) = \frac{1}{k+1}, \quad k = 0, \ldots, p-1, \quad (14)
$$

$$
b^\top B^{l+1} \left[ \alpha^{k-l-1} + \gamma \delta_{k-l-1,1} \right] = \frac{1}{k-l} b^\top B^l \alpha^{k-l}, \quad (15)
$$

where $k = 1, \ldots, p-1, l = 1, \ldots, k-2$. Consider the case $k - l = 2$. Then we get for the condition (15) a new condition

$$
b^\top B^k e = \frac{1}{2} b^\top B^{k-2} \alpha^2, \quad k = 2, \ldots, p-1,
$$

which can be found in the paper of Ostermann and Roche [14]. In the $k - l = m$ we get

$$
b^\top B^{l+1} \alpha^{m-1} = \frac{1}{m} b^\top B^l \alpha^m, \quad m \geq 3, l = 0, \ldots, p-m-1,
$$
For DIRK–methods we get the order conditions
\[ b^T \alpha^k = \frac{1}{k+1}, \quad k = 0, \ldots, p, \] (16)
\[ b^T A^{l+1} \alpha^{k-l-1} = \frac{1}{k-l} b^T \alpha^{k-l}, \quad k = 1, \ldots, p, l = 1, \ldots, k-2. \] (17)

The order conditions (17) are satisfied if the simplifying conditions \( C(q) \) are satisfied (\([4, 23]\), i.e.
\[ A\alpha^{k-1} = \alpha^k / k. \]
But DIRK–methods with a regular coefficient matrix \( A \) can only satisfy \( C(1) \). In [13] similar order conditions are derived.

### 3.2 The local error in the stiff case

We start our considerations with the local error
\[ \delta_\tau(t_{m+1}) = u_{m+1} - \varphi(t_{m+1}) \]
\[ = \varphi_m - \varphi_{m+1} + z b^T (I - zB)^{-1} [\varphi_m e - \Phi - \tau \gamma \dot{\varphi}_m] \]
\[ + \tau b^T (I - zB)^{-1} [\dot{\Phi} + \tau \gamma \ddot{\varphi}_m]. \]

In the stiff case we have \( \tau \to 0 \) and simultaneously \( z \to \infty \) but in the stiff case \( z \) tends to infinity if \( \tau \to 0 \). In this case we can expand the term \( (I - zB)^{-1} \) for large \( z \) as follows
\[ (I - zB)^{-1} = -(Bz)^{-1} - (Bz)^{-2} - \ldots. \]

Since we need a Taylor expansion for the two variables \( \tau \) and \( z \) we need the derivatives of \( (I - zB)^{-1} \). Let \( \tilde{z} = 1/z \). Then
\[ \left( \frac{I - B}{\tilde{z}} \right)^{-1} = -\frac{k!}{B^k}, \quad \text{for} \quad \tilde{z} \to 0, \]
if \( k \geq 1 \). Then the Taylor expansion of \( \delta_\tau(t_{m+1}) \) reads as
\[ \delta_\tau(t_{m+1}) = \varphi_m - \sum_{k=0}^{p} \varphi_m^{(k)} \frac{\tau^k}{k!} + O(\tau^{p+1}) \]
\[ - z \sum_{k=1}^{p+1} b^T \sum_{l=1}^{k} \binom{k}{l} l! B^{-l} \left[ \varphi_m e \delta_{k-l,0} - \alpha^{k-l} \varphi_m^{(k-l)} - \gamma \dot{\varphi}_m \delta_{k-l,1} \right] \frac{\tau^{k-l}}{k! z^l} \]
\[ - \tau \sum_{k=1}^{p} b^T \sum_{l=1}^{k} \binom{k}{l} l! B^{-l} \left[ \alpha^{k-l} \varphi_m^{(k-l+1)} + \gamma \ddot{\varphi}_m \delta_{k-l,1} \right] \frac{\tau^{k-l}}{k! z^l}. \]
In the second row we have for the case $k = l$ that
\[
\varphi_m e \rho^{k-l} \varphi_m^{(k-l)} - \gamma \varphi_m \rho^{k-l,1} = 0.
\]

In the second row we sum up from $l = 0$ to $k - 1$ and in the third row from $k = 2$ to $p + 1$. Then we get by using the definition of the binomial coefficient
\[
\delta (t_{m+1}) = - \sum_{k=1}^{p} \varphi_m^{(k)} \tau^k_k + O(\tau^{p+1})
\]
\[
+ \sum_{k=2}^{p+2} \sum_{l=0}^{k-2} b^T B^{-l-1} \left[ \alpha^{k-l-1} \varphi_m^{(k-l-1)} + \gamma \varphi_m \delta^{k-l-1,1} \right] \frac{\tau^{k-l-1}}{(k-l-1)! z^l}
\]
\[
- \sum_{k=2}^{p+1} \sum_{l=1}^{k-1} b^T B^{-l} \left[ \alpha^{k-l} \varphi_m^{(k-l)} + \gamma \varphi_m \delta^{k-l,1} \right] \frac{\tau^{k-l}}{(k-l-1)! z^l}.
\]

In the next step we split the sum in the second row
\[
\delta (t_{m+1}) = - \sum_{k=1}^{p} \varphi_m^{(k)} \tau^k_k + \sum_{k=2}^{p+1} b^T B^{-1} \left[ \alpha^{k-1} + \gamma \delta^{k-1,1} \right] \varphi_m^{(k-1)} \frac{\tau^{k-1}}{(k-1)!} + O(\tau^{p+1})
\]
\[
+ \sum_{k=3}^{p+2} \sum_{l=1}^{k-2} b^T B^{-l-1} \left[ \alpha^{k-l-1} + \gamma \delta^{k-l-1,1} \right] \varphi_m^{(k-l-1)} \frac{\tau^{k-l-1}}{(k-l-1)! z^l}
\]
\[
- \sum_{k=2}^{p+1} \sum_{l=1}^{k-1} b^T B^{-l} \left[ \alpha^{k-l} + \gamma \delta^{k-l-1,1} \right] \varphi_m^{(k-l)} \frac{\tau^{k-l}}{(k-l-1)! z^l}.
\]

Summing up in the second row from $k = 2$ to $p - 1$ and collecting some terms leads to
\[
\delta (t_{m+1}) = \sum_{k=2}^{p} \left[ b^T B^{-1} \alpha^k - 1 \right] \varphi_m^{(k)} \frac{\tau^k_k}{k!} + O(\tau^{p+1})
\]
\[
+ \sum_{k=2}^{p+1} \sum_{l=1}^{k-1} b^T B^{-l-1} \left[ \alpha^{k-l} + \gamma \delta^{k-l,1} \right] \varphi_m^{(k-l)} \frac{\tau^{k-l}}{(k-l)! z^l}
\]
\[
- \sum_{k=2}^{p+1} \sum_{l=1}^{k-1} b^T B^{-l} \left[ \alpha^{k-l} + \gamma \delta^{k-l,1} \right] \varphi_m^{(k-l)} \frac{\tau^{k-l}}{(k-l-1)! z^l}.
\]
and finally
\[ \delta_\tau(t_{m+1}) = \sum_{k=2}^{p-1} \left[ \mathbf{b}^\top \mathbf{B}^{-1} \mathbf{\alpha}^k - 1 \right] \varphi_m^{(k)} \frac{\tau^k}{k!} + \mathcal{O}(\tau^{p+1}) \]
\[ + \sum_{k=2}^{p+1} \mathbf{b}^\top \sum_{l=1}^{k-2} \left\{ \mathbf{B}^{-l-1} \left[ \mathbf{\alpha}^{k-l} + \mathbf{\gamma} \delta_{k-l,1} \right] \frac{1}{(k-l)} \right. \]
\[ - \mathbf{B}^{-l} \left[ \mathbf{\alpha}^{k-l-1} + \mathbf{\gamma} \delta_{k-l-1,1} \right] \right\} \varphi_m^{(k-l)} \frac{\tau^{k-l}}{(k-l-1)!} z^l, \]

since
\[ \mathbf{B}^{-1} \left[ \mathbf{\alpha}^{k-l} + \mathbf{\gamma} \delta_{k-l,1} \right] \frac{1}{(k-l)} - \left[ \mathbf{\alpha}^{k-l-1} + \mathbf{\gamma} \delta_{k-l-1,1} \right] = 0 \]

for \( l = k - 1 \). For Rosenbrock methods we get the order conditions
\[ \mathbf{b}^\top \mathbf{B}^{-1} \mathbf{\alpha}^k = 1, \quad k = 2, \ldots, p, \quad (18) \]
\[ \mathbf{b}^\top \mathbf{B}^{-(l+1)} \frac{1}{k-l} \mathbf{\alpha}^{k-l} = \mathbf{b}^\top \mathbf{B}^{-l} \left[ \mathbf{\alpha}^{k-l-1} + \mathbf{\gamma} \delta_{k-l-1,1} \right], \quad (19) \]

for \( l = 1, \ldots, k - 2 \) and \( k = 1, \ldots, p + 1 \). For the DIRK–methods we obtain
\[ \mathbf{b}^\top \mathbf{A}^{-1} \mathbf{c}^k = 1, \quad k = 2, \ldots, p, \quad (20) \]
\[ \mathbf{b}^\top \mathbf{A}^{-(l+1)} \frac{1}{k-l} \mathbf{c}^{k-l} = \mathbf{b}^\top \mathbf{A}^{-l} \mathbf{c}^{k-l-1}, \quad l = 1, \ldots, k - 1, k = 1, \ldots, p + 1. \quad (21) \]

4 New methods of order 2

4.1 A second order Rosenbrock–Wanner method

To obtain a 2nd order stiffly accurate ROW–method with three internal stages the order conditions
\[ \left\{ \begin{array}{l}
\mathbf{b}_1 + \mathbf{b}_2 + \mathbf{b}_3 = 1 \\
\mathbf{b}_2 \mathbf{\beta}_1' + \mathbf{b}_3 \mathbf{\beta}_3' = \frac{1}{2} - \mathbf{\gamma}
\end{array} \right. \quad (22) \]

should be satisfied. Here we use the abbreviations \( \mathbf{\beta}_{ij} := \mathbf{\alpha}_{ij} + \mathbf{\gamma}_{ij} \) and \( \mathbf{\beta}_i' := \sum_{j=1}^{i-1} \mathbf{\beta}_{ij} \). Moreover the embedded method should be stiffly accurate, too, i.e. we have \( \mathbf{\alpha}_{21} = 1, \mathbf{\beta}_i = \mathbf{\beta}_2, i = 1, 2, 3, \) and \( \mathbf{\beta}_{21} = 1 - \mathbf{\gamma} \). To obtain full order 2 the new order condition
\[ \mathbf{b}^\top \mathbf{B}^{-2} \mathbf{\alpha}^2 = 2 \quad (23) \]
should be satisfied to avoid order reduction. First we rewrite the order condition (PR2) as

\[ 2\gamma^3 - \gamma (b_2 \alpha_2^2 + b_3 \alpha_3^2) + 2b_3 \beta_32\alpha_2^2 = 0. \]

Inserting the conditions for a stiffly accurate method we get

\[ 2\gamma^2 - \gamma + \beta_{32} = 0. \tag{24} \]

The other conditions read as

\begin{align*}
\beta_{31} + \beta_{32} & = 1 - \gamma, \tag{25} \\
\beta_{32}\beta_{21} & = 1/2 - 2\gamma + \gamma^2, \tag{26} \\
\alpha_{21} & = \alpha_3 = 1, \tag{27} \\
\beta_{21} & = 1 - \gamma. \tag{28}
\end{align*}

First we compute \( \beta_{32} \) from (24) and (26). This leads to a cubic equation for determining \( \lambda \), i.e.

\[ 4\gamma^3 - 8\gamma^2 + 6\gamma - 1 = 0, \]

which has the real solution

\[ \gamma = \frac{-x^2 + 4x + 2}{6x}, \quad x = (17 + 3\sqrt{33})^{1/3}. \]

Then it follows

\begin{align*}
\beta_{32} & = \gamma - 2\gamma^2, \\
\beta_{21} & = \frac{1/2 - 2\gamma + \gamma^2}{\gamma - 2\gamma^2}, \\
\beta_{31} & = 1 - 2\gamma - 2\gamma^2.
\end{align*}

All coefficients of the new method called ROS2PR are presented in Table 1.

### 4.2 A second order DIRK–method

Our new DIRK–method should have 3 internal stages, be of order 2, and satisfy the order condition (21) for \( l = 1 \) and \( k = 3 \). Moreover the method and its embedded method should be stiffly accurate. In this case the conditions (20) are automatically satisfied. Then the conditions (16) for \( k = 1, 2, \)
Table 1: Set of coefficients for ROS2PR

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$2.2815549365396182e - 01$</th>
<th>$\alpha_{21}$</th>
<th>$1.00000000000000000e + 00$</th>
<th>$\gamma_{21}$</th>
<th>$-2.2815549365396182e - 01$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha_{31}$</td>
<td>$0.00000000000000000e + 00$</td>
<td>$\gamma_{31}$</td>
<td>$6.4779887126104239e - 01$</td>
<td>$\alpha_{32}$</td>
<td>$1.00000000000000000e + 00$</td>
</tr>
<tr>
<td>$\gamma_{32}$</td>
<td>$-8.7595436491500420e - 01$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_1$</td>
<td>$6.4779887126104239e - 01$</td>
<td>$\hat{b}_1$</td>
<td>$7.7184450634603818e - 01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_2$</td>
<td>$1.2404563508499580e - 01$</td>
<td>$\hat{b}_2$</td>
<td>$2.2815549365396182e - 01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$b_3$</td>
<td>$2.2815549365396182e - 01$</td>
<td>$\hat{b}_3$</td>
<td>$0.00000000000000000e + 00$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(21) for $l = 1$ and $k = 3$ and $a_{21} + a_{11} = 1$ should be satisfied. Moreover $\gamma := a_{ii}$, $i = 1, \ldots, 3$. Summarising these order conditions we have

\[
a_{31} + a_{32} = 1 - \gamma, \quad (29)
\]

\[
a_{32}a_{21} = \gamma^2 - 2\gamma + \frac{1}{2}, \quad (30)
\]

\[
a_{21} = 1 - \gamma, \quad (31)
\]

\[
a_{32}a_{21} = \gamma - 3\gamma^2 + \gamma^3. \quad (32)
\]

From condition (30) and (32) follows a condition for $\gamma$, i.e.

\[
\gamma^3 - 4\gamma^2 + 3\gamma - \frac{1}{2} = 0,
\]

which has a solution $\gamma \approx 0.2372$. Then we can formulate the other coefficients w.r.t. $\gamma$. We have

\[
a_{32} = \frac{\gamma^2 - 2\gamma + 1/2}{1 - \gamma}
\]

and

\[
a_{31} = 1 - \gamma - \frac{\gamma^2 - 2\gamma + 1/2}{1 - \gamma}.
\]

We call our new method DIRK2PR and summarise the coefficients in Table 2.

### 5 Numerical results

In this section we compare our new methods ROS2PR and DIRK2PR with other second order DIRK– and ROW–methods which are listed in Table 3.
Table 2: Set of coefficients for DIRK2PR

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$a_{21}$</th>
<th>$a_{31}$</th>
<th>$a_{32}$</th>
<th>$c_1$</th>
<th>$c_2$</th>
<th>$c_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2.3728621957824146e-01$</td>
<td>$7.6271378042175854e-01$</td>
<td>$6.5555390873299095e-01$</td>
<td>$1.0715987168876759e-01$</td>
<td>$2.3728621957824146e-01$</td>
<td>$1.0000000000000000e+00$</td>
<td>$0.0000000000000000e+00$</td>
</tr>
</tbody>
</table>

Table 3: Properties of the selected DIRK– and ROW–methods

<table>
<thead>
<tr>
<th>Name</th>
<th>s</th>
<th>p</th>
<th>Embedding</th>
<th>$R(\infty)$</th>
<th>stiffly acc.</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crank-Nicolson (CN)</td>
<td>2</td>
<td>2</td>
<td>no</td>
<td>1</td>
<td>yes</td>
<td>[4]</td>
</tr>
<tr>
<td>Ellsiepen</td>
<td>2</td>
<td>2</td>
<td>yes</td>
<td>0</td>
<td>yes</td>
<td>[2]</td>
</tr>
<tr>
<td>DIRK2PR</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>0</td>
<td>yes</td>
<td>see Section 4.2</td>
</tr>
<tr>
<td>ROS2</td>
<td>2</td>
<td>2</td>
<td>yes</td>
<td>0</td>
<td>no</td>
<td>[24]</td>
</tr>
<tr>
<td>Scholz4-5</td>
<td>2</td>
<td>2</td>
<td>no</td>
<td>-1</td>
<td>no</td>
<td>[21]</td>
</tr>
<tr>
<td>ROS2S</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>0</td>
<td>yes</td>
<td>[5]</td>
</tr>
<tr>
<td>ROS2PR</td>
<td>3</td>
<td>2</td>
<td>yes</td>
<td>0</td>
<td>yes</td>
<td>see Section 4.1</td>
</tr>
</tbody>
</table>

5.1 Example of Prothero–Robinson

First we consider the well-known example from Prothero and Robinson (4) with

$$\varphi(t) = \sin \left( \frac{\pi}{4} + t \right).$$

The ODE is solved (4) with equidistant step sizes $\tau = \frac{1}{10^2 \pi}, k = 0, \ldots, 13$ in the time interval $(0, 1/10]$. In Figure 1 we present the numerical results for $\lambda = -1$ (left) and $\lambda = -10^6$ (right). In the case $\lambda = -1$ all methods converge with order 2. The other case $\lambda = -10^6$ shows that the method of Ellsiepen has order reduction. The numerically observed order of convergence drops down to 1. But the convergence order of all other methods is 2. One interesting effect can be observed: The numerical error decreases down to $10^{-16}$ only for the methods ROS2S and Scholz 4-5. For all other methods the numerical
Figure 1: $\tau$ versus error for (4) with $\lambda = -1$ (left) and $\lambda = -10^6$ (right) error reduces only to $10^{-13}$. In the case of adaptive time step control we solve

the Prothero–Robinson problem until $\bar{t} = 100$ is reached. Moreover we made no simulations with the Scholz 4-5 method and the scheme of Crank-Nicolson since for these schemes no embedded methods are available (see [17] for the scheme of Crank-Nicolson). Similar results to the previous test case can be observed. In the case $\lambda = -1$ all methods behave well. In the stiff case, i.e. $\lambda = -10^6$, the method ROS2 produces the most inexact results. The method of Ellsiepen performs better than ROS2, but the best methods are DIRK2PR, ROS2PR and ROS2S.

5.2 Incompressible Navier–Stokes equations

Let $J$ be a time interval and $\Omega \subset \mathbb{R}^d$ be a domain. We consider the incompressible Navier–Stokes equations which are given in dimensionless form
by
\[
\dot{u} - Re^{-1}\Delta u + (u \cdot \nabla)u + \nabla p = f \quad \text{in} \ J \times \Omega, \\
\nabla \cdot u = 0 \quad \text{in} \ J \times \Omega, \\
u = g \quad \text{on} \ J \times \partial\Omega, \\
u(0, x) = u_0 \quad x \in \Omega,
\]
(33)

where \(Re\) denotes the positive Reynolds number. Details to the discretisation in space and time can be found for example in [9] and the references cited in there. In our first example of the incompressible Navier–Stokes equations the right-hand side \(f\), the initial condition \(u_0\) and the non-homogeneous Dirichlet boundary conditions are chosen such that
\[
u_1(t, x, y) = t^3y^2, \\
u_2(t, x, y) = t^2x, \\
p(t, x, y) = tx + y - (t + 1)/2
\]
is the solution of (33). Moreover we set \(Re = 1, \Omega = (0, 1)^2\), and solve the problem in the time interval \((0, 1/10]\). We use the \(Q_2/P_1^{\text{disc}}\) discretisation on a uniform mesh which consists of squares with an edge length \(h = 1/32\). Note that for any \(t\) the solution can be represented exactly by discrete functions. Hence, all occurring errors will result from the temporal discretisation. During the calculations we have to deal with 8,450 d.o.f. for the velocity and 3,072 d.o.f. for the pressure. As time steps we use \(\tau = \frac{1}{10 \cdot 2^k}, \ k = 0, \ldots, 7\). The numerical results are presented in Figure 3. Considering the velocity error it can be observed that all chosen schemes converge with order 2 as expected. A similar observation can be made for the pressure error. But in this case the schemes ROS2 and the scheme of Ellsiepen give the most inaccurate results. The best results are obtained with the scheme of Crank-Nicolson.

Figure 3: \(\tau\) versus error for (33) velocity \(u\) (left) and pressure \(p\) (right)
and the method ROS2S. Good results are obtained with the new methods DIRK2PR and ROS2PR.

The flow around a cylinder which will be considered was defined as a benchmark problem in [20] and studied numerically in detail in [7]. Figure 4 presents the flow domain. The right hand side of the Navier-Stokes equations (33) is $f = 0$, the final time is $\bar{t} = 8$ and the inflow and outflow boundary conditions are given by

$$u(t, 0, y) = u(t, 2.2, y) = 0.41^{-2} \sin(\pi t/8)(6y(0.41-y), 0) \text{ m s}^{-1}, \ 0 \leq y \leq 0.41.$$  

On all other boundaries, the no-slip condition $u = 0$ is prescribed. The Reynolds number of the flow, based on the mean inflow, the diameter of the cylinder and the prescribed viscosity $\nu = 10^{-3} \text{ m}^2 \text{ s}^{-1}$ is $0 \leq Re(t) \leq 100$.

The coarsest grid (level 0) is presented in Figure 5. All computations are carried out on level 4 of the spatial grid refinement resulting in 107,712 velocity d.o.f. and 39,936 pressure d.o.f.

The characteristic values of the flow are the drag coefficient $c_d(t)$ and the lift coefficient $c_l(t)$ at the cylinder. These coefficients can be computed by

$$c_d(t) = -20 [(\dot{u}, v_d) + (\nu \nabla u, \nabla v_d) + ((u \cdot \nabla)u, v_d) - (p, \nabla v_d)]$$
$$c_l(t) = -20 [(\dot{u}, v_l) + (\nu \nabla u, \nabla v_l) + ((u \cdot \nabla)u, v_l) - (p, \nabla v_l)]$$

for all functions $v_d, v_l \in (H^1(\Omega))^2$ where $(v_d)|_S = (1, 0)^T$ and $(v_l)|_S = (0, 1)^T$, $S$ being the boundary of the body, and $v_d$ and $v_l$ vanish on all other boundaries. Another benchmark value in [20] is the difference of the pressure between the front and the back at the cylinder at the final time.
$p(8, 0.15, 0.2) - p(8, 0.25, 0.2)$. Reference values for this difference and the maximal values of the drag and the lift coefficient are given in [6, 9]. In

![Figure 6: CPU-time versus error for (4): drag (left), lift (middle) and $\Delta p$ (right)](image)

this example we use an adaptive timestep control (see [9] for details). Only the method of Crank–Nicolson is applied without time-adaptivity since embedding is not possible for this method (see [17]). If we compare the DIRK methods DIRK2PR and the method of Ellsiepen DIRK2PR delivers more accurate results. ROS2 delivers quite good numerical results but the computing time is too long. The best results are obtained with the methods ROS2S, ROS2PR, and DIRK2PR. Of course the method DIRK2PR takes a longer computing time than Rosenbrock methods ROS2S and ROS2PR, but it is faster than the Rosenbrock method ROS2S. This is surprising since for every time step non-linear systems have to be solved in case of a DIRK–method.

**Conclusions and outlook**

In this note we analyse the numerical error of DIRK– and ROW–methods when they are applied on the Prothero–Robinson example. We observed that the methods may have order reduction if certain order conditions are not satisfied. We presented two new methods (ROS2PR and DIRK2PR) satisfying new order conditions, and apply these methods on two stiff problems: the Prothero–Robinson example and the incompressible Navier–Stokes equations. Both methods deliver good results, and the convergence order of 2 is reached for the pressure component of the Navier–Stokes equations.

In the next step a DIRK–method with $p = 3$ should be created which approximates the pressure with order 3, too. Moreover the theory should be applied on DIRK–methods with a singular coefficient matrix $A$, i.e. $a_{11} = 0$. Convergence results for these methods should be found.
References


<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-04</td>
<td>B. Sun, M. Lochau, P. Huhn, U. Goltz</td>
<td>Parameter Optimization of an Engine Control Unit using Genetic Algorithms</td>
</tr>
<tr>
<td>2009-07</td>
<td>M. Huhn, C. Knieke</td>
<td>Semantic Foundation and Validation of Live Activity Diagrams</td>
</tr>
<tr>
<td>2010-01</td>
<td>A. Litvinenko and H. G. Matthies</td>
<td>Sparse data formats and efficient numerical methods for uncertainties quantification in numerical aerodynamics</td>
</tr>
<tr>
<td>2010-02</td>
<td>D. Grunwald, M. Lochau, E. Börger, U. Goltz</td>
<td>An Abstract State Machine Model for the Generic Java Type System</td>
</tr>
<tr>
<td>2010-03</td>
<td>M. Krosche, R. Niekamp</td>
<td>Low-Rank Approximation in Spectral Stochastic Finite Element Method with Solution Space Adaption</td>
</tr>
<tr>
<td>2011-02</td>
<td>B. V. Rosić, A. Litvinenko, O. Pajonk, H. G. Matthies</td>
<td>Direct Bayesian update of polynomial chaos representations</td>
</tr>
<tr>
<td>2011-04</td>
<td>O. Pajonk, B.V. Rosić, A. Litvinenko, and H. G. Matthies</td>
<td>A Deterministic Filter for non-Gaussian Bayesian Estimation</td>
</tr>
<tr>
<td>2011-05</td>
<td>H. G. Matthies</td>
<td>A Hitchhiker’s Guide to Mathematical Notation and Definitions</td>
</tr>
</tbody>
</table>
On Causal Semantics of Petri Nets

Extended Version of Model-based Coverage-Driven Test Suite Generation for Software Product Lines

2011-08 W.-B. Pöttner, J. Morgenroth, S. Schildt, L. Wolf
An Empirical Performance Comparison of DTN Bundle Protocol Implementations

2011-09 H. G. Matthies, A. Litvinenko, O. Pajonk, B. V. Rosić and E. Zander
Parametric and Uncertainty Computations with Tensor Product Representations

Parameter Identification in a Probabilistic Setting

Efficient Analysis of High Dimensional Data in Tensor Formats

2011-12 S. Oster
A Semantic Preserving Feature Model to CSP Transformation

2012-01 O. Pajonk, B. V. Rosić and H. G. Matthies
Deterministic Linear Bayesian Updating of State and Model Parameters for a Chaotic Model

2012-02 B. V. Rosić and H. G. Matthies
Stochastic Plasticity - A Variational Inequality Formulation and Functional Approximation Approach I: The Linear Case

2012-03 J. Rang
An analysis of the Prothero-Robinson example for constructing new DIRK and ROW methods