Mathematica
Balkanica

New Series Vol. 19, 2005, Fasc. 1-2

On the Convergence of Linearized Implicit Runge-Kutta Methods and their Use in Parameter Optimization

Matthias Gerdts

A parameter optimization problem subject to mechanical multibody dynamics is solved by a multiple shooting method. The equations of motion are discretized by linearized implicit Runge-Kutta (LIRK) methods, which only require the solution of a linear equation system instead of a nonlinear one in each integration step. This allows to use fixed step-sizes during integration and often leads to a speed-up in the numerical solution of the parameter optimization problem when compared to BDF methods with step-size and order selection. The maximum attainable order of convergence for ordinary differential equations is 4 depending on the appropriate choice of an initial guess. The capability of the method is demonstrated at a discretized optimal control problem resulting from automobile test-driving.

AMS Subj. Classification: 65L06, 65L20, 90C90

Key Words: linearized implicit Runge-Kutta method, order of convergence, stability, parameter optimization

1. Introduction

The equations of motion of a mechanical multibody system (MBS) are given by the following differential equation

$$M(x(t),p)\dot{x}(t) = g(x(t),p) \tag{1}$$

for a fixed time interval $t \in [t_0, t_f]$, $t_0 < t_f$ with the state $x(t) \in \mathbb{R}^{n_x}$, parameter vector $p \in \mathbb{R}^{n_p}$, generalized forces $g(\cdot) \in \mathbb{R}^{n_x}$ and symmetric and positive definite mass matrix $M(\cdot) \in \mathbb{R}^{n_c \times n_c}$. Technical applications often lead to the following parameter optimization (PO) problem

$$\begin{array}{ll} \text{Minimize} & J[x,p] = \Phi(x(t_1),\ldots,x(t_N),p) \\ \text{subject to} & \text{equation (1)}, \\ & C_I(x(t_1),\ldots,x(t_N),p) \leq 0, \\ & C_E(x(t_1),\ldots,x(t_N),p) = 0, \end{array}$$

with fixed time points $t_i \in [t_0, t_f]$, $i = 1, \ldots, N$, inequality constraints $C_I(\cdot) \in \mathbb{R}^{n_{C_I}}$, and equality constraints $C_E(\cdot) \in \mathbb{R}^{n_{C_E}}$. Typical PO problems are

- parameter identification problems, where t_i, $i = 1, \ldots, N$ denote measure points and the objective function

$$\Phi(x(t_1), \ldots, x(t_N), p) := \frac{1}{2} \sum_{i=1}^{N} \|y_i - h(t_i, x(t_i), p)\|_{W_i}^2$$

  results from a maximum likelihood approach with measurements $y_i = h(t_i, x(t_i), p) + \varepsilon_i$, where $\varepsilon_i \sim N(0, W_i)$ are normally distributed measurement errors with covariance matrices W_i.

- discretized optimal control problems, where t_i, $i = 1, \ldots, N$ denote control grid points $t_0 = t_1 < t_2 < \ldots < t_N = t_f$ and p denotes the control discretization parameters, e.g. $p = (u_1, u_2, \ldots, u_N)^\top$ with $u_i \approx u(t_i)$. The objective function is often of Mayer type, that is $J[x, p] = \Phi(x(t_f), p)$. The constraints usually are given by discretized path constraints and boundary conditions, cf. [2].

The above PO problem is solved numerically by a direct shooting method similar to the method described in [2]. The direct shooting method is based on the application of a sequential quadratic programming (SQP) method to the PO problem. In each iteration of the SQP method the dynamical constraints (1) are solved by an appropriate numerical integration scheme and the constraints of the PO problem can be evaluated. Unless optimality and feasibility tolerances specified for the SQP method are satisfied, the SQP method updates the current iterate and performs another iteration. In this article we focus particularly on a numerical method for solving (1): Linearized implicit Runge-Kutta (LIRK) methods.

2. Convergence of Linearized Runge-Kutta Methods

Mechanical systems often lead to stiff initial value problems (IVP). Their numerical solution requires the usage of implicit Runge-Kutta (IRK) methods, since explicit Runge-Kutta methods are not suitable for stiff problems. For notational convenience the following discussion is restricted to autonomous IVP's of type

$$\dot{x}(t) = f(x(t)), \qquad x(t_0) = x_0. \tag{2}$$

Notice, that (1) is of type (2) with $f(x) = M^{-1}(x) \cdot g(x)$. The explicit dependence on the parameter p is omitted here, since it is not essential for the following convergence investigations. Please notice furthermore, that it is always possible to transform a non-autonomous differential equation into an equivalent autonomous equation, cf. [1], pp. 137-138. Furthermore, it can be proved similarly to the proof of Lemma 4.16, p. 138 in [1] that the below defined LIRK method is invariant under autonomization, if and only if $\sum_{i=1}^{s} b_i = 1$ and

$$c_i = \sum_{j=1}^{s} a_{ij}, \qquad i = 1, \dots, s \tag{3}$$

hold for the coefficients of the upcoming LIRK method. One step from t_n to $t_{n+1} = t_n + h$ of a IRK method, compare [4], applied to (2) is given by

$$\eta(t_{n+1}) := \eta(t_n) + h\Phi(\eta(t_n), h), \quad \Phi(\eta(t_n), h) := \sum_{i=1}^{s} b_i k_i(\eta(t_n), h) \tag{4}$$

where $\eta(t_{n+1}) \approx x(t_{n+1})$ and the stage derivatives $k_i(\eta(t_n), h)$, $i = 1, \dots, s$ are implicitly defined by the nonlinear equation system

$$F(k, \eta(t_n), h) := \begin{pmatrix} k_1 - f\left(\eta(t_n) + h\sum_{j=1}^{s} a_{1j}k_j\right) \\ \vdots \\ k_s - f\left(\eta(t_n) + h\sum_{j=1}^{s} a_{sj}k_j\right) \end{pmatrix} = 0 \tag{5}$$

for $k = (k_1, \dots, k_s)$, which is solved numerically by Newton's method. Notice, that Equation (5) can be easily adapted to (1) by multiplying row i with M evaluated at $\eta(t_n) + h\sum_{j=1}^{s} a_{ij}k_j$.

To circumvent the computational effort needed to solve (5), what is particularly important if the IVP has to be solved many times, e.g. during the iterative solution of the above PO problem, a new class of integration methods – LIRK methods – is defined by performing only one iteration of Newton's method. Hence, the LIRK method is defined by the linear equation system

$$F'_k(k^{(0)}, \eta(t_n), h) \cdot \left(k - k^{(0)}\right) + F(k^{(0)}, \eta(t_n), h) = 0. \tag{6}$$

Herein, F'_k denotes the partial derivative of F w.r.t. k and $k^{(0)} = (k_1^{(0)}, \dots, k_s^{(0)})$ is an initial guess to be specified below. A similar idea is used for the construction of ROW methods, cf. [4].

In the sequel we use the abbreviations $\eta_n := \eta(t_n)$, $f_n := f(\eta(t_n))$, and $A := (a_{ij})$. In addition, we will exploit the relationship (3) extensively. We are up to investigate the convergence properties of the LIRK method and have to specify $k^{(0)}$ in detail. In this article we choose $k^{(0)} = (k_1^{(0)}, \dots, k_s^{(0)})$ with

$$k_i^{(0)} := f(\eta_n), \qquad i = 1, \dots, s. \tag{7}$$

Notice, that it is also possible to choose $k^{(0)}$ differently. For instance, the choice $k_i^{(0)} = 0$ for all $i = 1, \dots, s$ seems to be simpler. Unfortunately, this choice has some drawbacks, e.g. the resulting LIRK method is not invariant under autonomization and it only allows to create methods with maximal order of convergence 2.

With $k^{(0)}$ according to (7) Equation (6) reduces to

$$(I - hB(\eta_n, h))\,(k - e \otimes f_n) + c(\eta_n, h) = 0, \tag{8}$$

with

$$B(\eta_n, h) := \begin{pmatrix} a^1 \otimes f'(\eta_n + hc_1 f_n) \\ \vdots \\ a^s \otimes f'(\eta_n + hcs f_n) \end{pmatrix}, c(\eta_n, h) := \begin{pmatrix} f_n - f(\eta_n + hc_1 f_n) \\ \vdots \\ f_n - f(\eta_n + hc_s f_n) \end{pmatrix}, \tag{9}$$

where a^i denotes the i^{th} row of A, $e = (1, \dots, 1)^\top \in \mathbb{R}^s$, and $\otimes$ is the Kronecker product. In its components (8) is equivalent with

$$k_i = h \sum_{j=1}^{s} a_{ij} f'(\eta_n + hc_i f_n) \cdot (k_j - f_n) + f(\eta_n + hc_i f_n), \quad i = 1, \dots, s. \tag{10}$$

The LIRK method defined by (4) and (8) is a one-step method. It generates a grid function $\eta_h : G_h \to \mathbb{R}^{n_x}$ where $G_h := \{t_0, t_1, \dots, t_N\}$ is an equidistant grid with time points $t_i := t_0 + ih$, $i = 0, 1, \dots, N$ and constant step length $h := (t_f - t_0)/N$. It is important to mention, that the following results can be extended to non-constant step-lengths as well. Let x denote the exact solution of (2). The global error is defined by

$$e_h := \max_{t \in G_h} \|x(t) - \eta(t)\|. \tag{11}$$

The method is said to be convergent, if $e_h \to 0$ for $h \to 0$. The order of convergence is p, if $e_h = \mathcal{O}(h^p)$ for $h \to 0$.

Let y denote the exact solution of the IVP $\dot{y}(t) = f(y(t))$, $y(\hat{t}) = \hat{x}$ and let $\eta(\hat{t} + h)$ denote the result of one step of the LIRK method with step size h starting at $(\hat{t}, \hat{x})$. Then, the local error is defined by $le_h(\hat{t}, \hat{x}) := \eta(\hat{t}+h) - y(\hat{t}+h)$. The method is said to be consistent, if $le_h(\hat{t}, \hat{x})/h \to 0$ holds for $h \to 0$ and all $(\hat{t}, \hat{x})$ in a neighborhood of the exact solution of (2). The consistency order is p, if $le_h(\hat{t}, \hat{x}) = \mathcal{O}(h^{p+1})$ holds for $h \to 0$ and all $(\hat{t}, \hat{x})$ in a neighborhood of the exact solution of (2).

For a general convergence proof we need stability and consistency, cf. [6]. A sufficient condition for stability is the local lipschitz-continuity of the increment function Φ w.r.t. η in the exact solution x of (2), i.e. there exist constants $\delta > 0$, $h_\delta > 0$, and $L_\delta > 0$ such that

$$\|\Phi(\eta_1, h) - \Phi(\eta_2, h)\| \leq L_\delta \|\eta_1 - \eta_2\| \tag{12}$$

holds for all η_j, $j = 1, 2$ with $\|\eta_j - x(t)\| \leq \delta$ and all $t \in [t_0, t_f]$ and all $0 < h \leq h_\delta$. We have the well-known convergence theorem for general one-step methods, which is proved in [1], pp. 128-129.

Theorem 2.1. *Let Φ be a locally lipschitz-continuous function w.r.t. η in the exact solution x of (2). Let the consistency order be p. Then the method converges with order p.*

First, we show the lipschitz-continuity of Φ.

Theorem 2.2. *Let f be twice continuously differentiable. Then, Φ in (4) with k from (8) is locally lipschitz-continuous in the exact solution x of (2).*

Proof. Since Φ is a linear combination of k_i, $i = 1, \ldots, s$ it suffices to show the local lipschitz-continuity of k. Since f is twice continuously differentiable, the functions B and c in (8) are continuously differentiable w.r.t. η and h. In addition, it holds $B(\eta_n, h) \to A \otimes f'_n$ for $h \to 0$. Hence, for sufficiently small h it holds $\|hB(\eta_n, h)\| < 1$, which implies the non-singularity of $I - hB(\eta_n, h)$ for sufficiently small h. By the implicit function theorem there exists a continuously differentiable function $k(\eta_n, h)$ satisfying (8). Application of the mean-value theorem for vector functions yields

$$\|k(\eta_1, h) - k(\eta_2, h)\| = \| \int_0^1 k'_\eta(\eta_1 + t(\eta_2 - \eta_1), h)(\eta_2 - \eta_1) dt \| \leq L_\delta \|\eta_2 - \eta_1\|,$$

where $L_\delta := \max_{\|\zeta - x(t)\| \leq \delta} \|k'_\eta(\zeta, h)\|$ exists for $\delta > 0$ and all $t \in [t_0, t_f]$ and all sufficiently small h. ■

It remains to determine the consistency order p in $le_h(t, x) = \eta(t + h) - x(t + h) = \mathcal{O}(h^{p+1})$. This is done by Taylor expansion of the exact and the approximate solution. Herein, the j^{th} derivative of f is a j-linear mapping with

$$f^{(j)}(x) \cdot (h_1, \ldots, h_j) = \sum_{i_1, \ldots, i_j = 1}^{n} \frac{\partial^j f(x)}{\partial x_{i_1} \cdots \partial x_{i_j}} h_{1,i_1} \cdots h_{j,i_j}.$$

The Taylor expansion of the exact solution is given by

$$x(t + h) = x + hf + \frac{h^2}{2!} f' f + \frac{h^3}{3!} \left(f''(f, f) + f' f' f \right)$$

$$
\begin{aligned}
&+\frac{h^4}{4!}\left(f'''(f,f,f)+3f''(f'f,f)+f'f''(f,f)+f'f'f'f\right)\\
&+\frac{h^5}{5!}\Big(f^{(4)}(f,f,f,f)+6f'''(f'f,f,f)+4f''(f''(f,f),f)\\
&\qquad +4f''(f'f'f,f)+3f''(f'f,f'f)+f'f'''(f,f,f)\\
&\qquad +3f'f''(f'f,f)+f'f'f''(f,f)+f'f'f'f'f\Big)+\mathcal{O}(h^6),
\end{aligned}
$$

where x is evaluated at t and f at $x(t)$, cf. [1], pp. 148,150.

Due to lack of space we omit the intermediate computation steps and only state the final Taylor expansion of the approximate solution, which was obtained by Taylor expansion of f and f' in (10), multiple recursive evaluation and exploitation of (3). It holds

$$
\begin{aligned}
\eta(t+h) \;=\;& x+h\sum_{i=1}^{s} b_i f + h^2\sum_{i=1}^{s} b_i c_i f'f + h^3\sum_{i=1}^{s} b_i\left(\sum_{j=1}^{s} a_{ij}c_j f'f'f + \frac{c_i^2}{2}f''(f,f)\right)\\
&+h^4\sum_{i=1}^{s} b_i\left(\sum_{j,l=1}^{s} a_{ij}a_{jl}c_l f'f'f'f + \sum_{j=1}^{s} a_{ij}\frac{c_j^2}{2}f'f''(f,f)\right.\\
&\qquad\left.+c_i\sum_{j=1}^{s} a_{ij}c_j f''(f'f,f)+\frac{c_i^3}{6}f'''(f,f,f)\right)\\
&+h^5\sum_{i=1}^{s} b_i\left(\sum_{j=1}^{s} a_{ij}\frac{c_j^3}{6}f'f'''(f,f,f)+c_i\sum_{j=1}^{s} a_{ij}\frac{c_j^2}{2}f''(f''(f,f),f)\right.\\
&\qquad+\frac{c_i^2}{2}\sum_{j=1}^{s} a_{ij}c_j f'''(f'f,f,f)+\frac{c_i^4}{24}f^{(4)}(f,f,f,f)\\
&\qquad+\sum_{j,l,m=1}^{s} a_{ij}a_{jl}a_{lm}c_m f'f'f'f'f+\sum_{j,l=1}^{s} a_{ij}a_{jl}\frac{c_l^2}{2}f'f'f''(f,f)\\
&\qquad\left.+\sum_{j,l=1}^{s} a_{ij}c_j a_{jl}c_l f'f''(f'f,f)+c_i\sum_{j,l=1}^{s} a_{ij}a_{jl}c_l f''(f'f'f,f)\right)+\mathcal{O}(h^6).
\end{aligned}
$$

A comparison with the Taylor expansion of the exact solution reveals, that the terms involving the derivative $f''(f'f,f'f)$ are missing in the Taylor expansion of the approximate solution. Hence, the maximal attainable consistency order is 4. Together with Theorem 2.1 and the local lipschitz-continuity of Φ (cf. Theorem 2.2) we proved the following theorem.

Theorem 2.3. *Let (3) hold. The LIRK method (10) is convergent of order 1, if* $\sum_{i=1}^{s} b_i = 1$ *holds. It is convergent of order 2, if in addition*

$\sum_{i=1}^{s} b_i c_i = \frac{1}{2}$ *holds. It is convergent of order 3, if in addition*

$$\sum_{i,j=1}^{s} b_i a_{ij} c_j = \frac{1}{6}, \qquad \sum_{i=1}^{s} b_i c_i^2 = \frac{1}{3}$$

hold. It is convergent of order 4, if in addition

$$\sum_{i,j,l=1}^{s} b_i a_{ij} a_{jl} c_l = \frac{1}{24}, \; \sum_{i,j=1}^{s} b_i a_{ij} c_j^2 = \frac{1}{12}, \; \sum_{i,j=1}^{s} b_i c_i a_{ij} c_j = \frac{1}{8}, \; \sum_{i=1}^{s} b_i c_i^3 = \frac{1}{4}.$$

hold. $p = 4$ *is the maximal attainable order of convergence.*

Remark The above conditions are the usual conditions known for general IRK methods. For linear differential equations, the LIRK method coincides with the ordinary nonlinear RK method (5). Hence, the LIRK method is A-stable if the nonlinear RK method is A-stable, since A-stability is defined for linear differential equations. Numerical tests for differential algebraic equations (DAE's) can be found in [3].

3. Example

We consider the double-lane-change manoeuvre, which is a common test in automobile industry. The double-lane-change manoeuvre can be formulated as an optimal control problem with free final time and state constraints, cf. [5]. Due to lack of space, we only give a short sketch of the equations of motion of the car model used for the simulation. The equations of motion of the so-called single-track car model are given by

$$\begin{aligned}
\dot{X} &= v\cos(\psi - \beta), \qquad \dot{Y} = v\sin(\psi - \beta), \\
\dot{v} &= \frac{1}{m}\left[(U_r - F_{Lx})\cos\beta + U_f\cos(\delta+\beta) - S_r\sin\beta - S_f\sin(\delta+\beta)\right], \\
\dot{\beta} &= w_z - \frac{1}{m\cdot v}\left[(U_r - F_{Lx})\sin\beta + U_f\sin(\delta+\beta) + S_r\cos\beta + S_f\cos(\delta+\beta)\right], \\
\dot{\psi} &= w_z, \qquad \dot{w}_z = \frac{1}{I_{zz}}\left[S_f\cdot l_v\cdot\cos\delta - S_r\cdot l_h\right], \qquad \dot{\delta} = w_\delta.
\end{aligned}$$

The state consists of the position (X, Y) of the car, the absolute velocity v, the side slip angle β, the yaw angle ψ, the yaw angle velocity w_z, and the steering angle δ. The driver controls the steering angle velocity w_δ. The right-hand side of the differential equation depends on the air resistance $F_{Lx} = F_{Lx}(v)$, the slip angles $\alpha_f = \alpha_f(\delta, w_z, v, \beta)$ and $\alpha_r = \alpha_r(w_z, v, \beta)$, the lateral tyre forces $S_f = S_f(\alpha_f)$ and $S_r = S_r(\alpha_r)$, the longitudinal tyre forces $U_f = U_f(F_B, F_{Rf})$ and $U_r = U_r(F_B, F_{Rr}, M_{wheel})$, the rolling resistances $F_{Rf} = F_{Rf}(v)$ and $F_{Rr} =$

$F_{Rr}(v)$, the braking force F_B, and the motor torque applied to the rear wheel $M_{wheel}(\phi, i)$. The latter depends on the accelerator pedal position ϕ and the gear i. F_B, ϕ, and i are possible additional control variables. Nevertheless, for the upcoming numerical tests, F_B, ϕ, and i are fixed. The quantities m, I_{zz}, l_v, and l_h are constants.

As already mentioned above, the double-lane-change manoeuvre can be formulated as a state-constrained optimal control problem with free final time. For its numerical solution a direct discretization method is applied. Herein, the control is approximated by a continuous and piecewise linear function, which only depends on a finite number of parameters, i.e. the function values at certain grid points. The discretization results in a parameter optimization problem PO introduced in Section 1. The PO is solved iteratively by a sequential quadratic programming (SQP) method, which requires to solve the differential equation within each iteration, cf. [2]. In the sequel a LIRK method with constant step size h and a BDF method with automatic step size and order selection are used to solve the differential equation numerically. Table 1 summarizes the CPU times for the numerical solution of the PO problem obtained for the third order linearized 2-stage RADAUIIA method, cf. [4], p. 74, and a standard BDF method (DASSL) with automatic step size and order selection.

Table 1: CPU times for the numerical solution of the discretized optimal control problem by the linearized 2-stage RADAUIIA method respectively the BDF method for different numbers N of control grid points.

N	CPU LIRK (in [s])	OBJ LIRK	CPU BDF (in [s])	Obj. BDF	rel. error Obj.	Speedup
26	2.50	7.718303	26.63	7.718305	0.00000026	9.4 %
51	8.15	7.787998	120.99	7.787981	0.00000218	6.7 %
101	18.02	7.806801	208.40	7.806798	0.00000038	8.6 %
201	21.24	7.819053	171.31	7.819052	0.00000013	12.4 %
251	198.34	7.817618	1691.15	7.817618	0.00000000	11.4 %
401	615.31	7.828956	4800.09	7.828956	0.00000000	12.8 %

The relative error in Table 1 denotes the relative error between the respective optimal objective function values of the discretized optimal control problems. The number N denotes the number of grid points in the formulation of PO in Section 1. The speedup is the ratio of columns 2 and 4. Table 1 shows that the LIRK method in average only needs 10 % CPU time when compared to the BDF method. A comparison of the accuracies of the respective solutions reveals that the quality of the LIRK solution is as good as the BDF solution.

Conclusion

The order of convergence of the LIRK method depends on the choice of the initial guess for one step of Newton's method. An appropriate choice allows to create methods up to order 4. The methods are easy to implement and show a very good computational performance within the numerical solution of parameter optimization problems. Numerical examples show that the use of LIRK methods reduces the computation time considerably compared to standard BDF integration schemes, while the accuracy remains the same.

References

[1] P. Deuflhard, F. Bornemann. *Scientific Computing with Ordinary Differential Equations.* Texts in Applied Mathematics, vol. **42**, Springer, New York, 2002.

[2] M. Gerdts. Direct Shooting Method for the Numerical Solution of Higher Index DAE Optimal Control Problems. *Journal of Optimization Theory and Applications*, **117**, 2003, 267-294.

[3] M. Gerdts. Parameter Optimization in Mechanical Multibody Systems and Linearized Runge-Kutta Methods. *accepted for publication in the proceedings of the ECMI 2002 conference held at Jurmala/Latvia.*

[4] E. Hairer, G. Wanner. *Solving ordinary differential equations II: Stiff and differential-algebraic problems.* Springer Series in Computational Mathematics, vol. **14**, Springer, Berlin Heidelberg New York, 1996.

[5] T. Moder. Optimale Steuerung eines KFZ im fahrdynamischen Grenzbereich. Diploma thesis, Mathematical Institute, Technical University of Munich, 1994.

[6] H. J. Setter. *Analysis of Discretization Methods for Ordinary Differential Equations.* Springer Tracts in Natural Philosophy, vol. **23**, Springer, Berlin Heidelberg New York, 1973.

Matthias Gerdts
Department of Mathematics
University of Bayreuth
95440 Bayreuth, Germany
Matthias.Gerdts@uni-bayreuth.de

Received 30.09.2003