Selection Strategies for Set-Valued Runge-Kutta Methods

Robert Baier

University of Bayreuth, Mathematical Institute, Applied Mathematics, D-95440 Bayreuth, Germany robert.baier@uni-bayreuth.de

Abstract. A general framework for proving an order of convergence for set-valued Runge Kutta methods is given in the case of linear differential inclusions, if the attainable set at a given time should be approximated. The set-valued method is interpreted as a (set-valued) quadrature method with disturbed values for the fundamental solution at the nodes of the quadrature method. If the precision of the quadrature method and the order of the disturbances fit together, then an overall order of convergence could be guaranteed. The results are applied to modified Euler method to emphasize the dependence on a suitable selection strategy (one strategy leads to an order breakdown).

Keywords: set-valued Runge-Kutta methods, linear differential inclusions, selection strategies, modified Euler

1 Introduction

In this article we restrict our attention to the case of linear differential inclusions (LDI). For motivations and connections to other areas, see e.g. [8].

Problem 1.1. Consider the linear differential inclusion (LDI)

$$x'(t) \in A(t)x(t) + B(t)U$$
 (f. a. e. $t \in I = [t_0, T]$), (1)

$$x(t_0) \in X_0 \tag{2}$$

with matrix functions $A: I \to \mathbb{R}^{n \times n}$, $B: I \to \mathbb{R}^{n \times m}$ and sets $X_0 \in \mathcal{C}(\mathbb{R}^n)$, $U \in \mathcal{C}(\mathbb{R}^m)$. Hereby, $\mathcal{C}(\mathbb{R}^n)$ denotes the set of nonempty, convex, compact subsets of \mathbb{R}^n .

The fundamental solution of the corresponding matrix differential equation

$$X'(t) = A(t)X(t) \quad (\textit{f. a. e. } t \in I \) \ ,$$

$$X(\tau) = I \ .$$

is denoted by $\Phi(\cdot, \tau)$ for $\tau \in I$, where $I \in \mathbb{R}^{n \times n}$ is the unit matrix.

Definition 1.2. The attainable set $\mathcal{R}(t, t_0, X_0)$ at a given time $t \in I$ for Problem 1.1 is defined as

$$\mathcal{R}(t, t_0, X_0) = \{x(t) \mid x : I \to \mathbb{R}^n \text{ is an absolutely continuous}$$

solution of $(1)-(2)\}$.

The Aumann integral introduced in [1] is an important tool for the following.

Definition 1.3. Consider a set-valued function $F: I \to \mathbb{R}^n$ with images in $\mathcal{C}(\mathbb{R}^n)$ which is measurable and integrably bounded (see [1]).

Then, Aumann's integral is defined as

$$\int_{t_0}^T F(t)dt := \left\{ \int_{t_0}^T f(t)dt \mid f(\cdot) \text{ is an integrable selection of } F(\cdot) \right\}.$$

It serves as a tool for reducing the approximation of the attainable set at time T to a problem of studying a set-valued quadrature method (see (7)).

Notation 1.4. The arithmetic operations of sets

$$\begin{split} \lambda \cdot C &:= \{ \, \lambda \cdot c \mid c \in C \, \} & (scalar \ multiple) \ , \\ C + D &:= \{ \, c + d \mid c \in C, \ d \in D \, \} & (Minkowski \ sum) \ , \\ A \cdot C &:= \{ \, Ac \mid c \in C \, \} & (image \ under \ a \ linear \ mapping) \end{split}$$

are defined as usual for $C, D \in \mathcal{C}(\mathbb{R}^n)$, $\lambda \in \mathbb{R}$, $A \in \mathbb{R}^{k \times n}$.

We denote with $d_H(C, D)$ the Hausdorff-distance of these two sets. The support function for C in direction $l \in \mathbb{R}^n$ is defined as

$$\delta^*(l,C) := \max_{c \in C} \langle l, c \rangle.$$

Lemma 1.5. Let $C, D \in \mathcal{C}(\mathbb{R}^n)$, $l \in \mathbb{R}^n$, $\lambda \geq 0$ and $A, B \in \mathbb{R}^{m \times n}$. Then,

$$\delta^*(l, C+D) = \delta^*(l, C) + \delta^*(l, D), \quad \delta^*(l, \lambda C) = \lambda \delta^*(l, C) \quad , \quad (3)$$

$$d_H(C, D) = \sup_{\|l\|_2 = 1} |\delta^*(l, C) - \delta^*(l, D)| , \qquad (4)$$

$$d_H(AU, BU) \le ||A - B|| \cdot ||U|| \text{ with } ||U|| := \sup_{u \in U} ||u||_2,$$
 (5)

$$d_H((A+B)U, AU + BU) \le ||A-B|| \cdot ||U|| . \tag{6}$$

In Problem 1.1, the attainable set at time T

$$\mathcal{R}(T, t_0, X_0) = \Phi(T, t_0) X_0 + \int_{t_0}^T \Phi(T, t) B(t) U dt$$
 (7)

could be rewritten as a sum of the transformed starting set and Aumann's integral of $\Phi(T,\cdot)B(\cdot)U$ (cf. e.g. [6]), where $\Phi(T,t)$ is the corresponding fundamental solution. Scalarization of (7) by support functions and applying the calculus rules in (3) and [4] yields for $l \in S_{n-1}$ (i.e., $l \in \mathbb{R}^n$ with $||l||_2 = 1$)

$$\delta^*(l, \mathcal{R}(T, t_0, X_0)) = \delta^*(l, \Phi(T, t_0)X_0) + \int_{t_0}^T \delta^*(l, \Phi(T, t)B(t)U)dt . \tag{8}$$

2 Quadrature and Combination Methods

Notation 2.1. For a given interval $I := [t_0, T]$ and a function $f : I \to \mathbb{R}^n$ consider the point-wise quadrature formula

$$Q(f; [t_0, T]) := \sum_{\mu=1}^{s} b_{\mu} f(t_0 + c_{\mu}(T - t_0))$$

for the approximation of $\int_I f(t)dt$, where $b_{\mu} \in \mathbb{R}$ are the weights and $c_{\mu} \in [0,1]$ determine the nodes $(\mu=1,\ldots,s)$. Introducing the step-size $h=\frac{T-t_0}{N}$ for $N \in \mathbb{N}$ and applying the quadrature formula on each sub-interval $[t_j,t_{j+1}]$ with $t_j=t_0+jh,\ j=0,\ldots,N-1$, we arrive at the iterated quadrature formula

$$Q_N(f;[t_0,T]) := h \sum_{j=0}^{N-1} Q(f;[t_j,t_{j+1}]) = h \sum_{j=0}^{N-1} \sum_{\mu=1}^{s} b_{\mu} f(t_j + c_{\mu} h) .$$

Definition 2.2. Consider a point-wise quadrature formula of Notation 2.1. Using the arithmetic operations of Notation 1.4, we introduce for a set-valued function $F: I \Rightarrow \mathbb{R}^n$ with images in $\mathcal{C}(\mathbb{R}^n)$ the iterated set-valued quadrature formula

$$Q_N(F;[t_0,T]) := h \sum_{j=0}^{N-1} \sum_{\mu=1}^{s} b_{\mu} F(t_j + c_{\mu} h).$$
(9)

These set-valued quadrature methods are studied by several authors, cf. e.g. [12,6,4,9,2]. Essential for reaching the same order of convergence as in the pointwise case is the smoothness of the function $t \mapsto \delta^*(l, F(t))$ uniformly in $l \in S_{n-1}$ due to the scalarization as in (8). To express the smoothness in a weaker sense, the averaged modulus of smoothness $\tau_k(f; h)$, presented e.g. in [11], is used which is a L_1 -norm of the local modulus of smoothness expressed as a certain supremum of the k-th finite difference of the function $f(\cdot)$.

Theorem 2.3. Let $F: I \Rightarrow \mathbb{R}^n$ with images in $\mathcal{C}(\mathbb{R}^n)$ be measurable and bounded. Consider a point-wise quadrature formula with precision p-1, $p \in \mathbb{N}$ (cf. [11]) and the set-valued iterated form (9) with step-size $h = \frac{T-t_0}{N}$, $N \in \mathbb{N}$. Then, the set-valued quadrature formula fulfills

$$d_H(\int_I F(t)dt, Q_N(F; I)) \le (1 + \sum_{\mu=1}^s \frac{b_\mu}{T - t_0}) \cdot W_p \cdot \sup_{\|l\|_2 = 1} \tau_p(\delta^*(l, F(\cdot)), \frac{2}{p}h) .$$

Proof. For the point-wise result see [2, Satz 1.2.11] which is based on [11, Theorem 3.4]. Apply this result for the function $t \mapsto \delta^*(l, F(t))$ for each $l \in S_{n-1}$ and use the equivalent expression (4) in Lemma 1.5 for the Hausdorff distance.

Set-valued quadrature methods could be used to approximate attainable set at the time T, if the values of the fundamental solution are known at the integration nodes $t_j + c_\mu h$, $\mu = 1, \dots, s$, $j = 0, \dots, N-1$. Otherwise, these values of

the fundamental solution needs to be approximated carefully so that the order of convergence of the quadrature method is not destroyed. Compare the next proposition with a result in [4] formulated with global disturbances.

Proposition 2.4. Let us consider Problem 1.1, set $h := \frac{T-t_0}{N}$, $N \in \mathbb{N}$ and the set-valued quadrature method with precision p-1, $p \in \mathbb{N}$, studied in Theorem 2.3 with $\tau_p(\delta^*(l, \Phi(T, \cdot)B(\cdot)U), h) \leq Ch^p$ uniformly in $l \in S_{n-1}$. For $j = 0, \ldots, N-1$ let the approximations $\widetilde{\Phi}(t_{j+1}, t_j)$ of the values of the fundamental solution resp. $\widetilde{U}_{\mu}(\cdot)$ of the images of $\Phi(t_{j+1}, \cdot)B(\cdot)U$ for $t = t_j + c_{\mu}h$ fulfill:

$$\widetilde{\Phi}(t_{j+1}, t_j) = \Phi(t_{j+1}, t_j) + \mathcal{O}(h^{p+1}) ,$$

$$d_H(\widetilde{U}_{\mu}(t_j + c_{\mu}h), \ \Phi(t_{j+1}, t_j + c_{\mu}h)B(t_j + c_{\mu}h)U) = \mathcal{O}(h^p) \quad (\mu = 1, \dots, s) .$$

 $[\mathcal{O}(h^q) \text{ is understood uniformly in } j \text{ and } \mu.]$ Then, the combination method

$$X_{j+1}^{N} = \widetilde{\Phi}(t_{j+1}, t_j) X_j^{N} + h \sum_{\mu=1}^{s} b_{\mu} \widetilde{U}_{\mu}(t_j + c_{\mu}h) \quad (j = 0, \dots, N-1) , \quad (10)$$

$$X_0^N \in \mathcal{C}(\mathbb{R}^n) \text{ with } d_H(X_0, X_0^N) = \mathcal{O}(h^p)$$
 (11)

defined above satisfies the global estimate

$$d_H(\mathcal{R}(T, t_0, X_0), X_N^N) = \mathcal{O}(h^p) . \tag{12}$$

Especially, if approximations of the values of the fundamental solution

$$\widetilde{\Phi}_{\mu}(t_{j+1}, t_j + c_{\mu}h) = \Phi(t_{j+1}, t_j + c_{\mu}h) + \mathcal{O}(h^p) \quad (\mu = 1, \dots, s) ,$$

then the estimation (12) above also holds with the following setting:

$$\widetilde{U}_{\mu}(t_j + c_{\mu}h) = \widetilde{\Phi}_{\mu}(t_{j+1}, t_j + c_{\mu}h)B(t_j + c_{\mu}h)U \quad (\mu = 1, \dots, s)$$
.

3 Set-Valued Runge-Kutta Methods

Explicit Runge-Kutta methods could be expressed by the Butcher array (cf. [5])

For a starting value $\eta_0^N \in X_0^N$ and $j = 0, \dots, N-1, \nu = 1, \dots, s$ let us define

$$\eta_{j+1}^{N} = \eta_{j}^{N} + h \sum_{\nu=1}^{s} b_{\nu} \xi_{j}^{(\nu)} , \qquad (13)$$

$$\xi_j^{(\nu)} = A(t_j + c_{\nu}h) \left(\eta_j^N + h \sum_{\mu=1}^{\nu-1} a_{\nu,\mu} \xi_j^{(\mu)} \right) + B(t_j + c_{\nu}h) u_j^{(\nu)} , \quad u_j^{(\nu)} \in U . \tag{14}$$

 X_{j+1}^N and X_0^N consist of all possible iterates η_{j+1}^N in (13)–(14) resp. chosen starting values η_0^N and form the set-valued Runge-Kutta method. Additional restrictions on the selections $u_j^{(\nu)} \in U$ for $\nu = 1, \ldots, s$ need to be imposed on each subinterval $I_j = [t_j, t_{j+1}]$ to increase the order of convergence in the set-valued case. These restrictions define different selection strategies.

Modified Euler Method

The modified Euler method resp. the method of Euler-Cauchy/Heun could be described by the Butcher array as

In [14], the method of Euler-Cauchy is discussed in detail with the result that for this method, one could use either constant selections or two free selections at each subinterval $[t_j, t_{j+1}]$ and reach order of convergence 2 under suitable smoothness conditions. In [13], the proofs are presented for the same method even in the case of strongly convex nonlinear differential inclusions.

Lemma 3.1. If we consider Problem 1.1, then the modified Euler method could be rewritten for the constant selection strategy " $u_j^{(1)} = u_j^{(2)}$ " as the combination method (10) of Proposition 2.4 with the iterated midpoint rule and

$$\begin{split} Q(\varPhi(t_{j+1},\cdot)B(\cdot)U;[t_{j},t_{j+1}]) &:= h\varPhi(t_{j+1},t_{j}+\frac{h}{2})B(t_{j}+\frac{h}{2})U \ , \\ \widetilde{\varPhi}(t_{j+1},t_{j}) &:= I + hA(t_{j}+\frac{h}{2}) + \frac{h^{2}}{2}A(t_{j}+\frac{h}{2})A(t_{j}) \ , \\ \widetilde{U}_{1}(t_{j}+\frac{h}{2}) &:= \Big(B(t_{j}+\frac{h}{2}) + \frac{h}{2}A(t_{j}+\frac{h}{2})B(t_{j})\Big)U \ . \end{split}$$

For two free selections $u_j^{(1)}, u_j^{(2)} \in U$ we have the iterated trapezoidal rule and

$$Q(\Phi(t_{j+1},\cdot)B(\cdot)U;I_j) := \frac{h}{2} \left(\Phi(t_{j+1},t_j)B(t_j)U + \Phi(t_{j+1},t_{j+1})B(t_{j+1})U \right) ,$$

$$\widetilde{\Phi}(t_{j+1},t_j) := I + hA(t_j + \frac{h}{2}) + \frac{h^2}{2}A(t_j + \frac{h}{2})A(t_j) ,$$

$$\widetilde{U}_1(t_j) := B(t_j + \frac{h}{2})U + hA(t_j + \frac{h}{2})B(t_j)U , \quad \widetilde{U}_2(t_{j+1}) := B(t_j + \frac{h}{2})U .$$

Proposition 3.2. Assume that $A'(\cdot)$ and $B(\cdot)$ are Lipschitz in Problem 1.1 and that $\delta^*(l, \Phi(T, \cdot)B(\cdot)U)$ is absolutely continuous with a L_1 -representative of the derivative with bounded variation uniformly in $l \in S_{n-1}$.

Then, the modified Euler method in Lemma 3.1 with p=2 in (11) and constant selection converges at least with order 2, whereas the modified Euler method with two independent selections converges at least with order 1.

Proof. Both quadrature methods have precision 1 (cf. [2], [12]), yielding order of convergence 2 in Theorem 2.3, if the disturbances would be of order $\mathcal{O}(h^2)$. For constant selections the result follows from Proposition 2.4 together with (5) and careful Taylor expansions in the estimations below:

$$\begin{split} \widetilde{\varPhi}(t_{j+1},t_{j}) &= \varPhi(t_{j+1},t_{j}) + \mathcal{O}(h^{3}), \\ \mathrm{d}_{\mathrm{H}}(\widetilde{U}_{1}(t_{j}+\frac{h}{2}), \ \Big(I+\frac{h}{2}A(t_{j}+\frac{h}{2})\Big)B(t_{j}+\frac{h}{2})U) &= \mathcal{O}(h^{2}) \ , \\ \mathrm{d}_{\mathrm{H}}(\Big(I+\frac{h}{2}A(t_{j}+\frac{h}{2})\Big)B(t_{j}+\frac{h}{2})U, \ \varPhi(t_{j+1},t_{j}+\frac{h}{2})B(t_{j}+\frac{h}{2})U) &= \mathcal{O}(h^{2}) \end{split}$$

In the case of two free selections, the reasoning is similar, but only accuracy $\mathcal{O}(h)$ is possible in general (due to (15) and (16)):

$$d_{H}(\widetilde{U}_{1}(t_{j}), (I + hA(t_{j}))B(t_{j})U) \leq d_{H}(B(t_{j} + \frac{h}{2})U, (I + hA(t_{j}))B(t_{j})U)$$

$$+ d_{H}(hA(t_{j} + \frac{h}{2})B(t_{j})U, \{0_{\mathbb{R}^{n}}\}) = \mathcal{O}(h) , \qquad (15)$$

$$d_{H}((I + hA(t_{j}))B(t_{j})U, \Phi(t_{j+1}, t_{j})B(t_{j})U) = \mathcal{O}(h^{2}) ,$$

$$d_{H}(\widetilde{U}_{2}(t_{j+1}), \Phi(t_{j+1}, t_{j+1})B(t_{j+1})U) = \mathcal{O}(h) . \qquad \Box$$

$$(16)$$

The assumptions in Proposition 3.2 could be weakened by demanding only the bounded variation of $A'(\cdot)$ and $B(\cdot)$. Clearly, for the strategy with two free selections, only $A(\cdot)$ needs to be Lipschitz, $B(\cdot)$ should be bounded and $\delta^*(l, \Phi(T, \cdot)B(\cdot)U)$ be of bounded variation uniformly in $l \in S_{n-1}$.

Since in general (even for the time-independent case, compare also (6)),

$$(B + \frac{h}{2}AB)U \neq (BU + \frac{h}{2}ABU) = \frac{1}{2}(BU + hABU + BU) , \qquad (17)$$

$$d_{\mathrm{H}}\left((B + \frac{h}{2}AB)U, \frac{1}{2}\left(BU + (BU + hABU)\right)\right) = \mathcal{O}(h) , \qquad (18)$$

both selection strategies for modified Euler differ. The proof of (18) uses a similar trick as in (15). This phenomena is also observed in the context of discretization by Runge-Kutta methods of nonlinear optimal control problems in [7]. In this work, additional assumptions on the coercitivity (not fulfilled in Problem 1.1) and on the smoothness of the optimal control leads to the accuracy up to $\mathcal{O}(h^2)$ for state and control variables using different proof ideas.

Scalarization as in (8) or direct methods for optimal control problems in [3] lead to numerical implementations of both selection strategies. For the scalarization approach, support functions of left-hand and right-hand sides of the equation (10) are calculated. This leads to an iterative method (cf. [2] for more details), if one restricts the computation of the support functions (or points) to a finite number of normed directions $l^{(j)} \in \mathbb{R}^n$, $j = 1, \ldots, M$.

Example 3.3. (cf. [4]) Let
$$n=2$$
, $m=1$, $I=[0,1]$, set $A(t)=\begin{pmatrix}0&1\\0&0\end{pmatrix}$, $B(t)=\begin{pmatrix}0\\1\end{pmatrix}$ and $U=[-1,1]$. Since (17) is fulfilled here, both selection strategies for

modified Euler differ (cf. Figure 1). In Figure 1, the reference set (the combi-

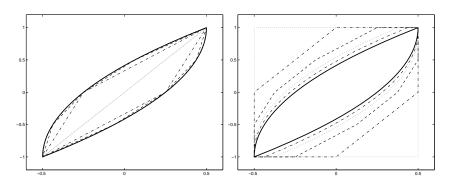


Fig. 1. modified Euler with constant (left picture) resp. 2 free selections (right one) (step sizes h=1,0.5,0.25,0.125,0.0625)

nation method "iterated trapezoidal rule and Euler/Cauchy" with N=10000 in [2]) is plotted with supporting points in M=200 directions with a thicker solid line, whereas the result for each calculated step size is depicted with dotted (h=1,0.125), dashed-dotted (h=0.5,0.0625) and dashed lines (h=0.25). One may recognize the different speed of convergence (2 resp. 1) even by the picture. This is underlined by the computed estimations of the order of convergence in Table 1. Hence, the possible order breakdown to $\mathcal{O}(h)$ in Proposition 3.2 for modified Euler with two free selections can occur for certain examples.

Table 1. convergence estimation of modified Euler for both selection strategies

	Hausdorff distance	estimated order	Hausdorff distance	estimated order
N	to reference set	of convergence	to reference set	of convergence
1	0.21434524		0.75039466	
2	0.05730861	1.90311	0.36454336	1.04156
4	0.01517382	1.91717	0.17953522	1.02182
8	0.00384698	1.97979	0.08841414	1.02192
16	0.00096510	1.99498	0.04419417	1.00042
	(constant selections)		(2 free selections)	

4 Conclusions

The presented framework may give a structural outline to proofs for linear differential inclusions with possibly non-optimal order of convergence for a set-valued

Runge-Kutta method with a chosen selection strategy. For the modified Euler method, the better selection strategy is formed by the constant selections which fits to the underlying (set-valued) quadrature method (i.e., the midpoint rule).

For the (classical) Runge-Kutta method of order 4 with Butcher array

first experiments show that the selection strategy with three free selections $u_j^{(\nu)}$, $\nu=1,2,4$ and $u_j^{(2)}=u_j^{(3)}$ lead to a set-valued method of at least order 3 under sufficient smoothness conditions. This fits best to the Simpson's rule as the underlying set-valued quadrature method. All presented selection strategies can be carried over to the case of nonlinear differential inclusions.

References

- R. J. Aumann. Integrals of Set-Valued Functions. J. Math. Anal. Appl., 12(1):1– 12, 1965.
- R. Baier. Mengenwertige Integration und die diskrete Approximation erreichbarer Mengen. Bayreuth. Math. Schr., 50:xxii + 248 S., 1995.
- R. Baier, C. Büskens, I. A. Chahma, and M. Gerdts. Approximation of Reachable Sets by Direct Solution Methods of Optimal Control Problems. submitted, 04/2004. 23 pages.
- R. Baier and F. Lempio. Computing Aumann's integral. In [10], pages 71–92, 1994.
- J. C. Butcher. The Numerical Analysis of Ordinary Differential Equations. John Wiley & Sons, Chichester-New York-Brisbane-Toronto-Singapore, 1987.
- T. D. Donchev and E. Farkhi. Moduli of smoothness of vector valued functions of a real variable and applications. *Numer. Funct. Anal. Optim.*, 11(5 & 6):497–509, 1990.
- A. L. Dontchev, W. W. Hager, and V. M. Veliov. Second-Order Runge-Kutta Approximations in Control Constrained Optimal Control. SIAM J. Numer. Anal., 38(1):202–226, 2000.
- 8. A. L. Dontchev and F. Lempio. Difference methods for differential inclusions: A survey. SIAM Rev., 34(2):263–294, 1992.
- M. Krastanov and N. Kirov. Dynamic interactive system for analysis of linear differential inclusions. In [10], pages 123–130, 1994.
- A. B. Kurzhanski and V. M. Veliov, editors. Modeling Techniques for Uncertain Systems, Proceedings of a Conferences held in Sopron, Hungary, July 6-10, 1992, volume 18 of Progress in Systems and Control Theory, Basel, 1994. Birkhäuser.
- B. Sendov and V. Popov. Averaged Moduli of Smoothness. Applications in Numerical Methods and Approximation. John Wiley and Sons, Chichester-New York-Brisbane-Toronto-Singapore, 1988.

- 12. V. M. Veliov. Discrete approximations of integrals of multivalued mappings. C. R. Acad. Bulgare Sci., 42(12):51-54, 1989.
- 13. V. M. Veliov. Second order discrete approximations to strongly convex differential inclusions. *Systems Control Lett.*, 13(3):263–269, 1989.
- 14. V. M. Veliov. Second Order Discrete Approximation to Linear Differential Inclusions. SIAM J. Numer. Anal., 29(2):439–451, 1992.