

Explicit stabilized integrators for stiff optimal control problems

Ibrahim Almuslimani

Joint work with

Gilles Vilmart (Geneva)

Inria-Rennes
Université de Rennes 1

Sample days
Brest September 2021

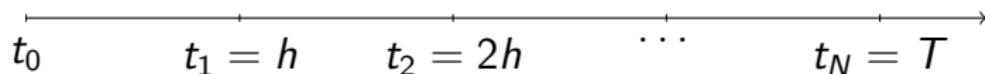
contents

- 1 Explicit stabilized methods
- 2 Preliminaries on optimal control
- 3 Second order Runge-Kutta-Chebyshev method for optimal control problems
- 4 Numerical experiments

I. Almuslimani and G. Vilmart. Explicit stabilized integrators for stiff optimal control problems. SIAM J. Sci. Comput., 43(2):A721–A743, 2021.

Stability analysis of deterministic integrators

$$\dot{y} := \frac{dy(t)}{dt} = f(y(t)), \quad y(0) = y_0, \quad t \in [0, T].$$



We denote by y_k the approximation of $y(t_k)$.

Stability analysis of deterministic integrators

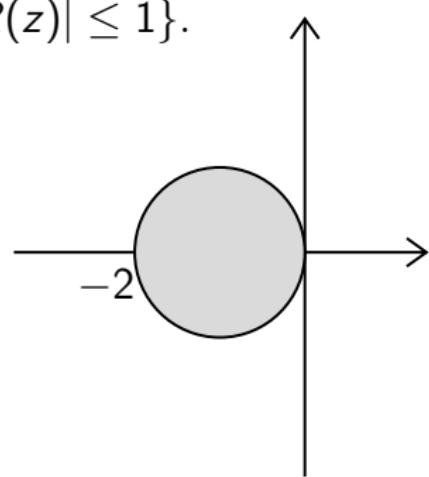
Dahlquist's linear test problem: $\dot{y}(t) = \lambda y$, $y(0) = y_0$.

Motivated by dissipative problems, we consider $\Re(\lambda) < 0$.

Stability function. A Runge-Kutta method with stepsize h yields $y_{k+1} = R(h\lambda)y_k$ and by induction we get $y_k = R(h\lambda)^k y_0$.

Stability domain $\mathcal{S} := \{z \in \mathbb{C}; |R(z)| \leq 1\}$.

Example. Explicit Euler method: $R(z) = 1 + z$,
the stability condition is $|1 + h\lambda| \leq 1$ and for
real eigenvalues $-2 \leq h\lambda \leq 0$ which leads for
diffusion problems $h \leq C\Delta x^2$ (severe stepsize
restriction).



Explicit stabilized methods (1960)

In order to construct an explicit stabilized integrator of order p , the first step is to find a polynomial $R_s(z)$ of degree s and order p , i.e.

$$R_s(z) = 1 + z + \cdots + \frac{z^p}{p!} + \mathcal{O}(z^{p+1}),$$

that solves the following problem

$$\text{Find } R_s(z) = 1 + z + \frac{z^2}{2!} + \cdots + \frac{z^p}{p!} + \alpha_{p+1}z^{p+1} + \cdots + \alpha_sz^s,$$

$$|R_s(z)| \leq 1 \text{ for } z \in [-\ell_s^p, 0], \text{ with } \ell_s^p \text{ as large as possible.}$$

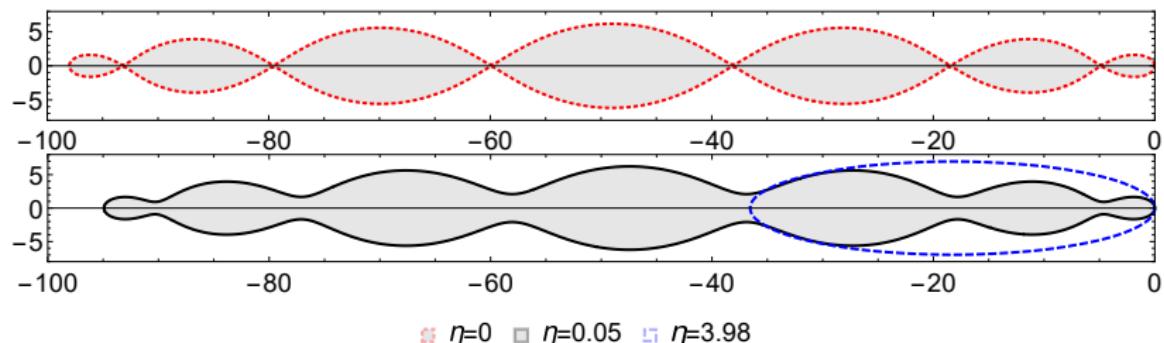
Deterministic Chebyshev method

$$\text{Solution for } p = 1 : \quad T_s\left(1 + \frac{z}{s^2}\right), \quad \ell_s^p = 2s^2.$$

Damping:

$$R_{s,\eta}(z)y_0 = \frac{T_s(\omega_0 + \omega_1 z)}{T_s(\omega_0)}, \quad \omega_0 = 1 + \frac{\eta}{s^2}, \quad \omega_1 = \frac{T_s(\omega_0)}{T'_s(\omega_0)},$$

$$\ell_s^p \geq (2 - 4/3\eta)s^2.$$



First order Deterministic Chebyshev method [1960]

$$\begin{aligned}y_{k0} &= y_k, \quad K_1 = y_{k0} + h\mu_1 f(y_{k0}), \\y_{ki} &= \mu_i h f(y_{k,i-1}) + \nu_i y_{k,i-1} + \kappa_i y_{k,i-2}, \quad i = 2, \dots, s \\y_{k+1} &= y_{ks},\end{aligned}$$

where, $k = 0, \dots, N$, $\mu_1 = \frac{\omega_1}{\omega_0}$, and for all $i = 2, \dots, s$

$$\mu_i = \frac{2\omega_1 T_{i-1}(\omega_0)}{T_i(\omega_0)}, \quad \nu_i = \frac{2\omega_0 T_{i-1}(\omega_0)}{T_i(\omega_0)}, \quad \kappa_i = 1 - \nu_i.$$

The number of stages s is chosen adaptively such that

$$\lambda_{\max} h \leq \ell_s^p s^2,$$

typically,

$$s := \left[\sqrt{\frac{h\lambda_{\max} + 1.5}{\ell_s^p}} + 0.5 \right].$$

Second order methods

To design a second order method, we need the stability polynomial to satisfy

$$R(z) = 1 + z + \frac{z^2}{2} + \mathcal{O}(z^3).$$

Correction introduced by Bekker in 1971:

$$R_s^\eta(z) = a_s + b_s T_s(\omega_0 + \omega_2 z),$$

where,

$$a_s = 1 - b_s T_s(\omega_0), \quad b_s = \frac{T_s''(\omega_0)}{(T_s'(\omega_0)^2)}, \quad \omega_0 = 1 + \frac{\eta}{s^2},$$

$$\omega_2 = \frac{T_s'(\omega_0)}{T_s''(\omega_0)}, \quad \eta = 0.15.$$

$\ell_s^p \simeq 0.65s^2$, optimal for second order: $\ell_s^p \simeq 0.82$ [Abdulle 2001].

Second order RKC method [1980]

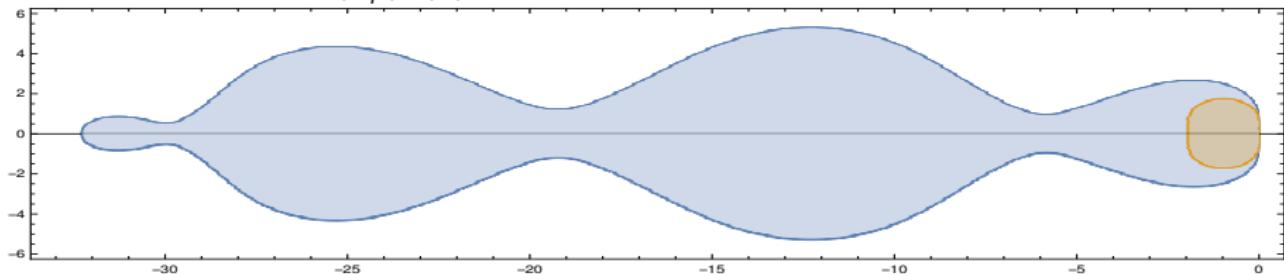
$$y_{k0} = y_k,$$

$$y_{k1} = y_{k0} + hb_1\omega_2 f(y_{k0}),$$

$$\begin{aligned} y_{ki} &= y_{k0} + \mu'_i h(f(y_{k,i-1}) - a_{i-1} f(y_{k0})) + \nu'_i (y_{k,i-1} - y_{k0}) \\ &\quad + \kappa'_i (y_{k,i-2} - y_{k0}), \end{aligned}$$

$$y_{k+1} = y_{ks},$$

where $k = 0, \dots, N - 1$, and for $i = 2, \dots, s$, $\mu'_i = \frac{2b_i\omega_2}{b_{i-1}}$, $\nu'_i = \frac{2b_i\omega_0}{b_{i-1}}$,
 $\kappa'_i = -\frac{b_i}{b_{i-2}}$, $b_i = \frac{T''_i(\omega_0)}{(T'_i(\omega_0)^2)}$, $a_i = 1 - b_i T_i(\omega_0)$.



Stability

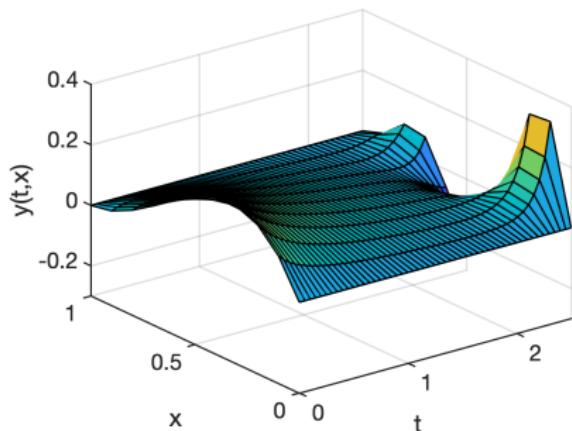


Figure: Heun method

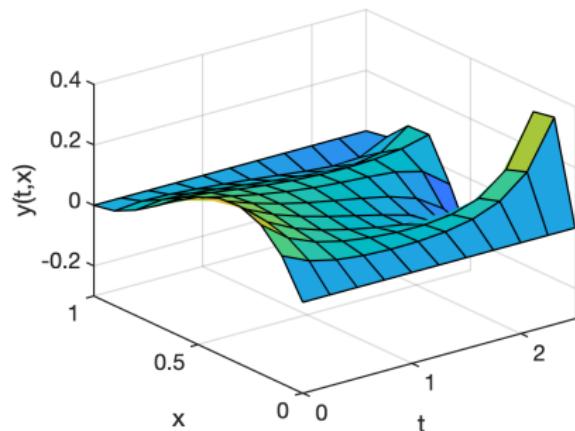


Figure: Explicit stabilized method

Explicit stabilized integrators (Chebyshev methods)

Yuan'Chzao Din 58', Franklin 59', Guillou, Lago, Saul'ev 60' ...

- RKC: Methods based on three-term recurrence relation (non-optimal) with $\ell_s^P \simeq 0.66 \cdot s^2$
van der Houwen, Shampine, Sommeijer, Verwer (RKC, IMEY extension IRKC, 1980-2007), Zbinden (PRKC 2011)
- Methods based on composition (no-recurrence relation)
Bogatyrev, Lebedev, Skvorstov, Medovikov (DUMKA 1976-2004), Jeltsch, Torrilhon 2007
- ROCK methods (close to optimal stability for second order)
Abdulle, Medovikov (ROCK2 2000-02) with $\ell_s^P \simeq 0.81 \cdot s^2$
Abdulle (ROCK4 2002-05) with $\ell_s^P \simeq 0.35 \cdot s^2$
- Extension to stiff stochastic problems: S-ROCK methods
Weak order 1: Abdulle, Cirilli, Li, Hu (S-ROCK 2007-2009, τ -ROCK methods 2010) with $\ell_s^P \simeq 0.33 \cdot s^2$
Abdulle, A., Vilmart (SK-ROCK methods 2018) with optimal stability domain $I_s^P \simeq 2.s^2$, and order 2 for sampling the invariant measure of ergodic SDEs.
Weak order 2: Abdulle, Vilmart, Zygalakis (S-ROCK2 2013) with $\ell_s^P \simeq 0.43 \cdot s^2$

- 1 Explicit stabilized methods
- 2 Preliminaries on optimal control
- 3 Second order Runge-Kutta-Chebyshev method for optimal control problems
- 4 Numerical experiments

Optimal control problem

We consider the following optimal control problem

$$\begin{aligned} & \min_u \Psi(y(T)); \\ & \dot{y}(t) = f(u(t), y(t)), \quad t \in [0, T], \\ & y(0) = y^0, \end{aligned} \tag{P}$$

with given final time $T > 0$ and initial condition $y^0 \in \mathbb{R}^d$.

- $y : [0, T] \rightarrow \mathbb{R}^d$ is the unknown state function,
- $u : [0, T] \rightarrow \mathbb{R}^m$ is the unknown control function.

If we discretize (P) using a RK discretization ($b_i > 0$ for all i) we naturally get the following discrete optimization problem,

$\min \Psi(y_N); \quad \text{subject to:}$

$$y_{k+1} = y_k + h \sum_{i=1}^s b_i f(u_{ki}, y_{ki}), \quad (\text{DP})$$

$$y_{ki} = y_k + h \sum_{j=1}^s a_{ij} f(u_{kj}, y_{kj}).$$

Pontryagin's maximum (or minimum) principle on (P) with $H(u, y, p) := p^T f(u, y)$:

$$\begin{aligned} \dot{y}(t) &= f(u(t), y(t)) = \nabla_p H(u(t), y(t), p(t)), \\ \dot{p}(t) &= -\nabla_y f(u(t), y(t))p = -\nabla_y H(u(t), y(t), p(t)), \\ 0 &= \nabla_u H(u(t), y(t), p(t)), \\ t &\in [0, T], \quad y(0) = y^0, \quad p(T) = \nabla \Psi(y(T)). \end{aligned} \quad (\text{OC})$$

Introducing **Lagrange multipliers** for the finite dimensional optimization problem (DP):

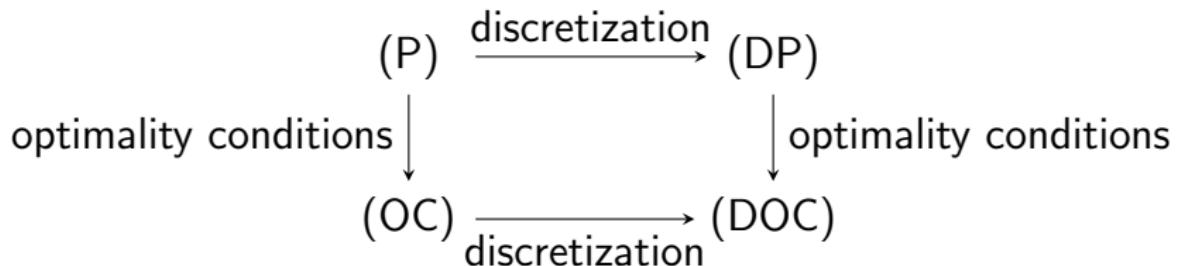
$$\begin{aligned}y_{k+1} &= y_k + h \sum_i b_i f(u_{ki}, y_{ki}), \quad y_{ki} = y_k + h \sum_j a_{ij} f(u_{kj}, y_{kj}), \\p_{k+1} &= p_k - h \sum_i \hat{b}_i \nabla_y H, \quad p_{ki} = p_k - h \sum_j \hat{a}_{ij} \nabla_y H, \\0 &= \nabla_u H(u_{kj}, y_{kj}, p_{kj}), \quad y_0 = y^0, \quad p_N = \nabla \Psi(y_N),\end{aligned}\tag{DOC}$$

where $k = 0, \dots, N-1$, $i = 1, \dots, s$, and the coefficients \hat{b}_i and \hat{a}_{ij} are defined by the following relations which correspond to the **symplecticity** conditions of partitioned RK methods,

$$b_i = \hat{b}_i, \quad \hat{a}_{ij} := b_j - \frac{b_j}{b_i} a_{ji}, \quad i, j = 1, \dots, s.$$

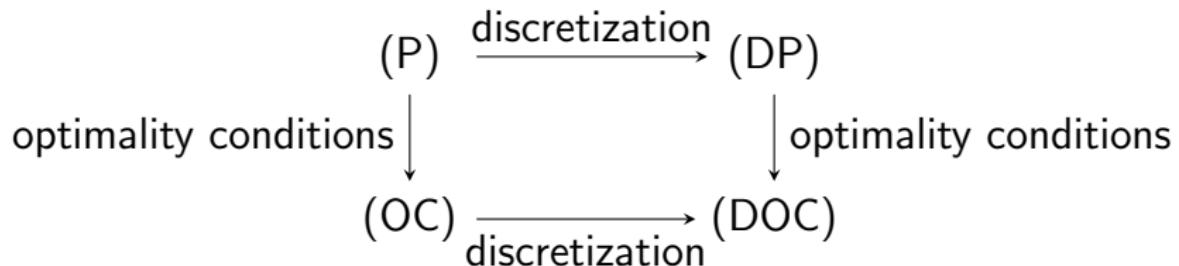
$$\hat{b}_i = b_i, \quad b_i \hat{a}_{ij} + \hat{b}_j a_{ji} = b_i \hat{b}_j, \quad i, j = 1, \dots, s. \quad (*)$$

For a RK method the following diagram commutes when $(*)$ is satisfied [Hager '00, Bonnans & Laurent-Varin '06]:



$$\hat{b}_i = b_i, \quad b_i \hat{a}_{ij} + \hat{b}_j a_{ji} = b_i \hat{b}_j, \quad i, j = 1, \dots, s. \quad (*)$$

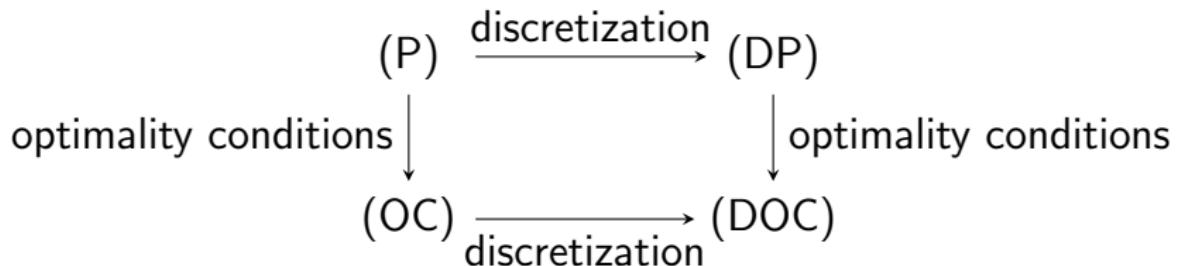
For a RK method the following diagram commutes when $(*)$ is satisfied [Hager '00, Bonnans & Laurent-Varin '06]:



- (a_{ij}, b_i) of order 2 with $(*) \implies$ order 2 for optimal control.

$$\hat{b}_i = b_i, \quad b_i \hat{a}_{ij} + \hat{b}_j a_{ji} = b_i \hat{b}_j, \quad i, j = 1, \dots, s. \quad (*)$$

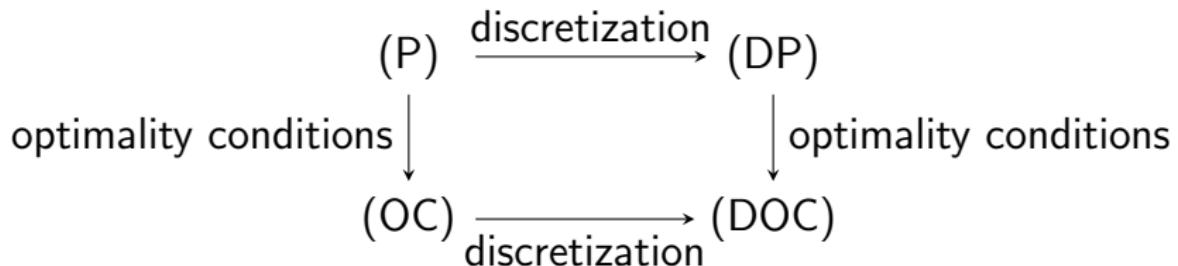
For a RK method the following diagram commutes when $(*)$ is satisfied [Hager '00, Bonnans & Laurent-Varin '06]:



- (a_{ij}, b_i) of order 2 with $(*) \implies$ order 2 for optimal control.
- (a_{ij}, b_i) of order 2 without $(*) \implies$ only order 1 for optimal control (**order reduction**) in general.

$$\hat{b}_i = b_i, \quad b_i \hat{a}_{ij} + \hat{b}_j a_{ji} = b_i \hat{b}_j, \quad i, j = 1, \dots, s. \quad (*)$$

For a RK method the following diagram commutes when $(*)$ is satisfied [Hager '00, Bonnans & Laurent-Varin '06]:



- (a_{ij}, b_i) of order 2 with $(*) \implies$ order 2 for optimal control.
- (a_{ij}, b_i) of order 2 without $(*) \implies$ only order 1 for optimal control (**order reduction**) in general.
- (a_{ij}, b_i) of order > 2 with (or without) $(*) \implies$ only order 2 for optimal control (**order reduction**) in general.

$$p_{k+1} = p_k - h \sum_i \hat{b}_i \nabla_y H, \quad p_{ki} = p_k - h \sum_j \hat{a}_{ij} \nabla_y H, \quad p_N = \nabla \Psi(y_N).$$

A calculation yields

$$p_N = \nabla \Psi(y_N), \quad \text{for } K = N-1, \dots, 0, \quad i = s, \dots, 1,$$

$$p_k = p_{k+1} + h \sum_{i=1}^s \tilde{b}_i \nabla_y H, \quad \tilde{b}_i = b_i,$$

$$p_{ki} = p_{k+1} + h \sum_{j=1}^s \tilde{a}_{ij} \nabla_y H, \quad \tilde{a}_{ij} = \frac{b_j}{b_i} a_{ji}.$$

- $(\tilde{a}_{ji}, \tilde{b}_i)$ is the **time adjoint** (Φ_{-h}^{-1}) of $(\hat{a}_{ji}, \hat{b}_i)$.
- We call $(\tilde{a}_{ji}, \tilde{b}_i)$ the **double adjoint** of (a_{ji}, b_i) .

Theorem (A., Vilmart)

A RK method (a_{ij}, b_i) and its double adjoint $(\tilde{a}_{ji}, \tilde{b}_i)$ share the same stability function $R(z)$.

- 1 Explicit stabilized methods
- 2 Preliminaries on optimal control
- 3 Second order Runge-Kutta-Chebyshev method for optimal control problems
- 4 Numerical experiments

Second order Runge-Kutta-Chebyshev method for optimal control problems

Second order RKC [Van der Houwen and Sommeijer in 1980]:

$\min \Psi(y_N)$, such that

$$y_{k0} = y_k,$$

$$y_{k1} = y_{k0} + \mu_1 h f(u_{k0}, y_{k0}),$$

$$y_{ki} = \mu_i h f(u_{k,i-1}, y_{k,i-1}) + \nu_i y_{k,i-1} + (1 - \nu_i) y_{k,i-2}, \quad i = 2, \dots, s,$$

$$y_{k+1} = a_s y_{k0} + b_s T_s(\omega_0) y_{ks},$$

with stability function $a_s + b_s T_s(\omega_0 + \omega_2 z)$. The stability region of the above method contains the interval $[-0.65s^2, 0]$.

Theorem (A., Vilmart)

The double adjoint ($p_{k+1} \mapsto p_k$) of the above RKC scheme is given for $k = N - 1, \dots, 0$ by the following recurrence relations

$$p_N = \Psi'(y_N), \quad p_{k_s} = p_{k+1}, \quad p_{k,s-1} = p_{k_s} + \frac{\mu_s}{\nu_s} h H_y(u_{k,s-1}, y_{k,s-1}, p_{k_s}),$$

for $j = 2, \dots, s - 1$,

$$\begin{aligned} p_{k,s-j} &= \frac{\mu_{s-j+1}\alpha_{s-j+1}}{\alpha_{s-j}} h H_y(u_{k,s-j}, y_{k,s-j}, p_{k,s-j+1}) \\ &\quad + \frac{\nu_{s-j+1}\alpha_{s-j+1}}{\alpha_{s-j}} p_{k,s-j+1} + \frac{(1 - \nu_{s-j+2})\alpha_{s-j+2}}{\alpha_{s-j}} p_{k,s-j+2}, \end{aligned}$$

$$p_{k0} = \mu_1 \alpha_1 h H_y(u_{k0}, y_{k0}, p_{k1}) + \alpha_1 p_{k1} + (1 - \nu_2) \alpha_2 p_{k2} + a_s p_{k+1},$$

$$p_k = p_{k0},$$

$$H_u(u_{k,s-j}, y_{k,s-j}, p_{k,s-j+1}) = 0, \quad j = 1, \dots, s.$$

where the coefficients α_j are defined using the induction

$$\alpha_s = b_s T_s(\omega_0),$$

$$\alpha_{s-1} = \nu_s \alpha_s,$$

$$\alpha_{s-j} = \nu_{s-j+1} \alpha_{s-j+1} + (1 - \nu_{s-j+2}) \alpha_{s-j+2}, \quad j = 2 \dots s-1.$$

Idea of the proof:

- Calculate the Lagrangian of the discrete optimal control problem.

where the coefficients α_j are defined using the induction

$$\alpha_s = b_s T_s(\omega_0),$$

$$\alpha_{s-1} = \nu_s \alpha_s,$$

$$\alpha_{s-j} = \nu_{s-j+1} \alpha_{s-j+1} + (1 - \nu_{s-j+2}) \alpha_{s-j+2}, \quad j = 2 \dots s-1.$$

Idea of the proof:

- Calculate the Lagrangian of the discrete optimal control problem.
- Calculate the optimality conditions given by $\nabla \mathcal{L} = 0$.

where the coefficients α_j are defined using the induction

$$\alpha_s = b_s T_s(\omega_0),$$

$$\alpha_{s-1} = \nu_s \alpha_s,$$

$$\alpha_{s-j} = \nu_{s-j+1} \alpha_{s-j+1} + (1 - \nu_{s-j+2}) \alpha_{s-j+2}, \quad j = 2 \dots s-1.$$

Idea of the proof:

- Calculate the Lagrangian of the discrete optimal control problem.
- Calculate the optimality conditions given by $\nabla \mathcal{L} = 0$.
- **Rescale** the stages of the obtained double adjoint in order to make them of form $p_{k+1} + \mathcal{O}(h)$.

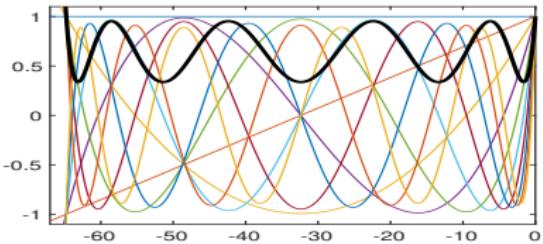


Figure: RKC.

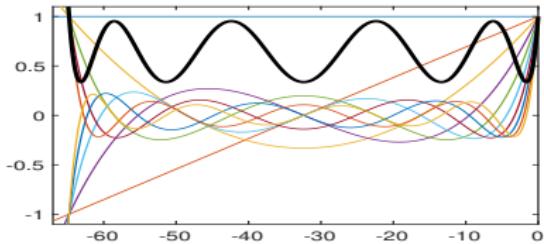


Figure: RKC double adjoint.

Theorem (A., Vilmart)

For $\eta = 0$, the stability functions of the internal stages of the RKC double adjoint satisfy $|R_{s,i}(z)| \leq 1$ for all $z \in [-\frac{2}{3}s^2 + \frac{2}{3}, 0]$.

Corollary

The above theorem is true for small damping $\eta > 0$.

Theorem (A., Vilmart)

The new method has order 2 for optimal control.

Difficulty for ROCK2

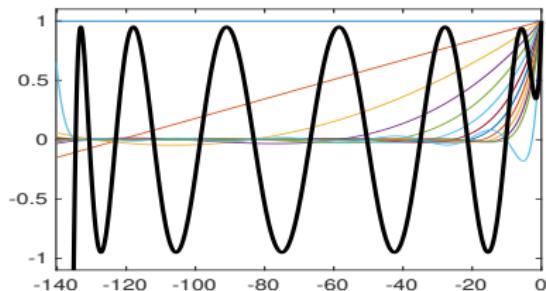


Figure: Classical ROCK2.

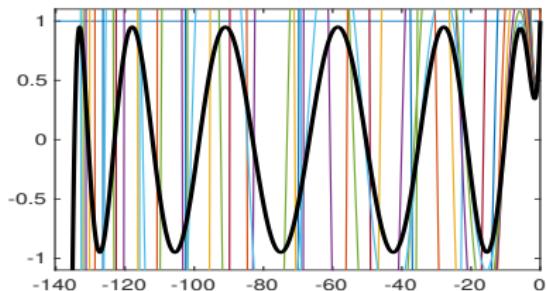


Figure: Double adjoint of ROCK2.

Figure: Internal stages (thin curves) and stability polynomials (bold curves) of the ROCK2 method and its double adjoint for optimal control for $s = 13$ stages.

$$\text{ROCK2: } R_s(z) = \color{red}w_2(z)\color{black} P_{s-2}(z).$$

$$\text{Double adjoint of ROCK2: } R_s(z) = P_{s-2}(z) \color{red}w_2(z).$$

Naively applying a RK method

Applying directly a (non-partitioned) RK method to:

$$\begin{aligned}\dot{y}(t) &= f(u(t), y(t)) = \nabla_p H(u(t), y(t), p(t)), \\ \dot{p}(t) &= -\nabla_y f(u(t), y(t))p = -\nabla_y H(u(t), y(t), p(t)), \\ 0 &= \nabla_u H(u(t), y(t), p(t)), \\ t &\in [0, T], \quad y(0) = y^0, \quad p(T) = \nabla \Psi(y(T)),\end{aligned}\tag{OC}$$

(with a shooting method) leads for diffusion problems

$$\partial_t p = -\Delta p.$$

- 1 Explicit stabilized methods
- 2 Preliminaries on optimal control
- 3 Second order Runge-Kutta-Chebyshev method for optimal control problems
- 4 Numerical experiments

Optimal control of a nonlinear diffusion-advection PDE (Burgers equation)

$$\min_{u \in L^2([0, T]; L^2(\Omega))} J(u) = \frac{1}{2} \|y(T) - y^{target}\|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \int_0^T \|u(t)\|_{L^2(\Omega)}^2$$

subject to

$$\begin{aligned} \partial_t y &= \mu \Delta y - \frac{\nu}{2} \partial_x(y^2) + u && \text{in } (0, T) \times \Omega, \\ y(0, x) &= g(x) && \text{in } \Omega, \\ y(t, x) &= 0 && \text{on } \partial\Omega. \end{aligned}$$

where $\mu = 0.1$, $\nu = 0.02$, in dimension $d = 1$ with domain $\Omega = (0, 1)$ and the final time is given by $T = 2.5$.

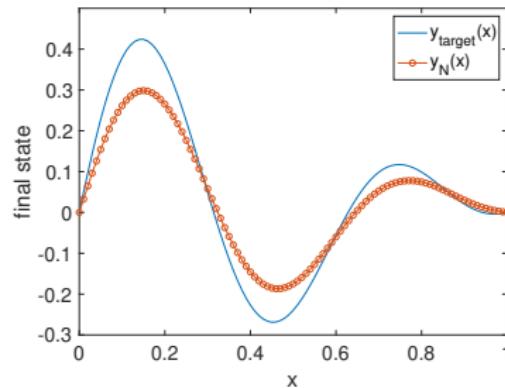
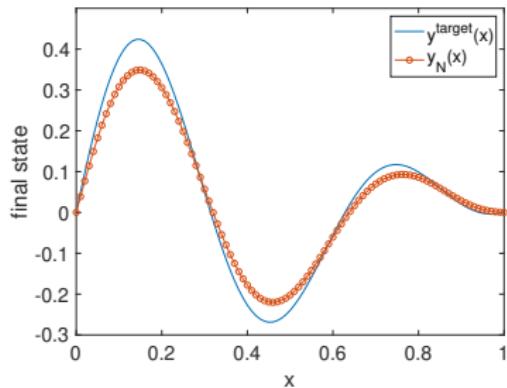
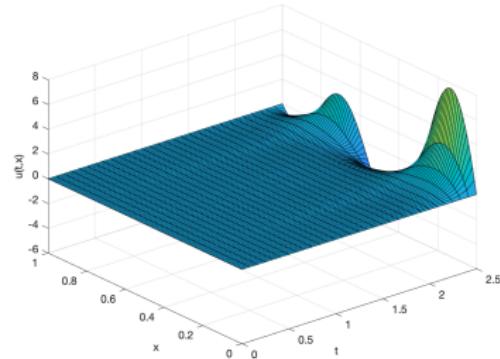
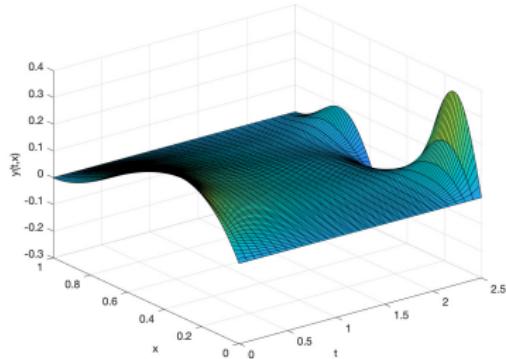


Figure: State, final state, and control. $\Delta x = 1/100$, $\Delta t = T/30$, $s = 24$ stages, and $\alpha = 0.01$. Down right: $\alpha = 0.02$.

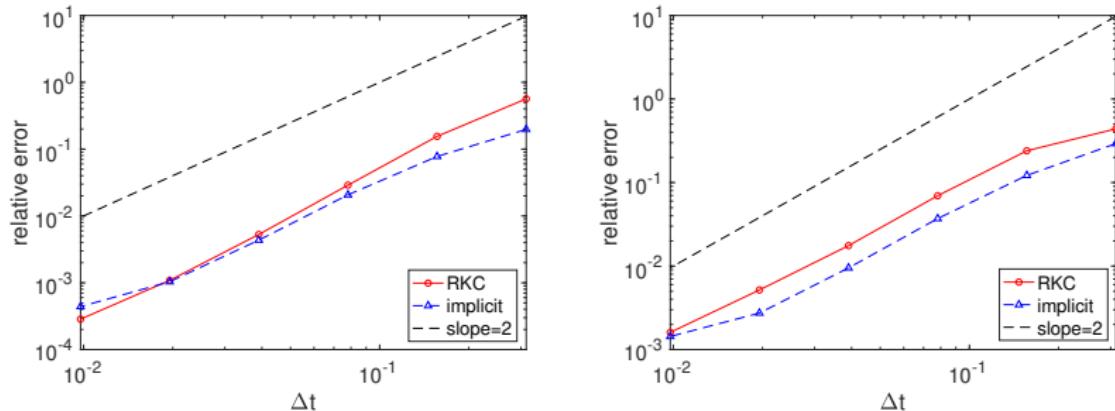


Figure: Error in the state (left) and the control (right) for $\Delta x = 10^{-2}$.

$$\begin{array}{c|cc} & \gamma & \\ \hline & 1 - 2\gamma & \gamma \\ \hline & 1/2 & 1/2 \end{array}$$

[Herty, Pareschi, and Steffensen (2013)] (IMEX)

Explicit stabilized methods for SDEs

Joint work with Assyr Abdulle (EPFL) and Gilles Vilmart (Geneva)

$$dX(t) = f(X(t))dt + \sum_{r=1}^m g^r(X(t))dW_r(t), \quad X(0) = X_0$$

$$K_0 = X_0$$

$$K_1 = X_0 + \mu_1 h f(X_0 + \nu_1 \sum_{r=1}^m g^r(X_0) \Delta W_j) + \kappa_1 \sum_{r=1}^m g^r(X_0) \Delta W_j$$

$$K_i = \mu_i h f(K_{i-1}) + \nu_i K_{i-1} + \kappa_i K_{i-2}, \quad i = 2, \dots, s.$$

$$X_1 = K_s,$$

$$R(p, q, \xi) = \frac{T_s(\omega_0 + \omega_1 p)}{T_s(\omega_0)} + \frac{U_{s-1}(\omega_0 + \omega_1 p)}{U_{s-1}(\omega_0)} \left(1 + \frac{\omega_1}{2} p\right) q \xi.$$

$$\mathbb{E}(|R(p, q, \xi)|^2) = \frac{T_s(\omega_0 + \omega_1 p)^2}{T_s(\omega_0)^2} + \frac{U_{s-1}(\omega_0 + \omega_1 p)^2}{U_{s-1}(\omega_0)^2} \left(1 + \frac{\omega_1}{2} p\right)^2 q^2.$$

Conclusion: We have used the new concept of the **double adjoint** of a Runge-Kutta method in the context of optimal control as a tool to construct **the first symplectic second order explicit stabilized methods** for unconstrained optimal control problems.

Perspectives:

- Constrained optimal control problems.
- Stochastic control problems.
- Oscillatory constraint equation (Uniformly accurate methods).

I. Almuslimani and G. Vilmart. Explicit stabilized integrators for stiff optimal control problems. SIAM J. Sci. Comput., 43(2):A721–A743, 2021.

Related work:

A. Abdulle, I. Almuslimani, and G. Vilmart. Optimal explicit stabilized integrator of weak order 1 for stiff and ergodic stochastic differential equations. SIAM/ASA J. Uncertain. Quantif., 6(2):937–964, 2018.