

# Processed splitting methods and high order actions in path integral Monte Carlo simulations

Fernando Casas

Fernando.Casas@uji.es

www.gicas.uji.es

Departament de Matemàtiques

Institut de Matemàtiques i Aplicacions de Castelló (IMAC)

Universitat Jaume I

Castellón, Spain

ICDSDEA 2010, Dresden

26 May, 2010

# Numerical integration of ODEs

- Given the ODE

$$x' = F(x), \quad x_0 = x(t_0) \in \mathbb{R}^D, \quad (1)$$

a one-step numerical *integrator*  $\psi_h$  for a time step  $h$  is of order  $\geq r$  if the numerical approximation

$$x_h \equiv \psi_h(x_0)$$

to the exact solution

$$x(h) = \exp(hF)(x_0)$$

is such that

$$x_h = x(h) + \mathcal{O}(h^{r+1})$$

# Processing in numerical integration

- First considered by **J. Butcher** in 1969 in the context of Runge–Kutta methods.
- Idea: to ‘enhance’ the integrator  $\psi_h$  by a change of variables given by a map  $\pi_h$  so that

$$\hat{\psi}_h = \pi_h \psi_h \pi_h^{-1} \quad (2)$$

is of order higher than  $\psi_h$ .

- $\psi_h$  is the **kernel** and  $\pi_h$  is the (post-)processor or corrector of the *processed* method  $\hat{\psi}_h$ .
- After  $n$  steps:  $\hat{\psi}_h^n = \pi_h \psi_h^n \pi_h^{-1}$
- *Pre-processor*  $\pi_h^{-1}$  applied only once, then  $\psi_h$  acts once per step and finally  $\pi_h$  is evaluated only when output is required

# Processing

- Advantageous if  $\hat{\psi}_h$  is more accurate than  $\psi_h$  and the cost of  $\pi_h$  is negligible
- Since the mid-1990s, it has been used in **geometric numerical integration** (Sanz-Serna *et al.*, McLachlan, Wisdom, Suzuki, Blanes, Murua, Ros, C., etc.).
- The method  $\psi_h$  is of *effective order*  $r$  if a post-processor  $\pi_h$  exists for which  $\hat{\psi}_h$  is of (conventional) order  $r$ , that is,

$$\pi_h \psi_h \pi_h^{-1} = \exp(hF) + \mathcal{O}(h^{r+1}).$$

# An example

- Suppose

$$x' = F(x) = F^{[a]}(x) + F^{[b]}(x),$$

and the  $h$ -flows  $\exp(hF^{[a]})$  and  $\exp(hF^{[b]})$  can be computed exactly.

- Then

$$\chi_h = e^{hF^{[b]}} e^{hF^{[a]}} \quad (3)$$

is a first order approximation,

$$\chi_h(x_0) = x(h) = e^{hF}(x_0) + \mathcal{O}(h^2)$$

- The same for its *adjoint*  $\chi_h^* = e^{hF^{[a]}} e^{hF^{[b]}}$

# An example

- The composition

$$\mathcal{S}_h^{[2]} = e^{\frac{h}{2}F[a]} e^{hF[b]} e^{\frac{h}{2}F[a]} \quad (4)$$

is second-order (leapfrog, Strang splitting, Störmer–Verlet method)

- ... But it *is* a processed method:

$$\begin{aligned} \mathcal{S}_h^{[2]} &= e^{\frac{h}{2}F[a]} e^{hF[b]} e^{\frac{h}{2}F[a]} = e^{\frac{h}{2}F[a]} e^{hF[b]} e^{hF[a]} e^{-hF[a]} e^{\frac{h}{2}F[a]} \\ &= e^{\frac{h}{2}F[a]} \chi_h e^{-\frac{h}{2}F[a]} = \pi_h \chi_h \pi_h^{-1} \end{aligned}$$

with  $\pi_h = \exp(\frac{h}{2}F[a])$

- Hence,  $\chi_h = e^{hF[b]} e^{hF[a]}$  is of effective order 2.

## Processing for splitting methods

- Given  $x' = F^{[a]}(x) + F^{[b]}(x)$ , leapfrog is an example of splitting method
- More general splitting methods:

$$\psi_h = e^{ha_1 F^{[a]}} e^{hb_1 F^{[b]}} e^{ha_2 F^{[a]}} \dots e^{ha_p F^{[a]}} e^{hb_p F^{[b]}} \quad (5)$$

with appropriately chosen coefficients  $a_i, b_i$  such that the scheme is of order  $r$ .

- The scheme is *symmetric* if  $a_i = a_{p+1-i}$ ,  $b_p = 0$  and  $b_i = b_{p-i}$ .
- By convention, method (5) has  $p$  stages
- We consider processed methods whose kernel is of the form (5)

# Advantages

- Order conditions: polynomial equations to be satisfied by the parameters of the method so that approximates the exact solution up to  $\mathcal{O}(h)^r$ .
- For a processed method  $\hat{\psi}_h$  of order  $r$ , many of the order conditions can be satisfied by using  $\pi_h$ .
- Consequence:  $\psi_h = e^{ha_1 F^{[a]}} e^{hb_1 F^{[b]}} e^{ha_2 F^{[a]}} \dots e^{ha_p F^{[a]}} e^{hb_p F^{[b]}}$  must fulfill a much reduced set of constraints, and it is possible to have kernels of effective order  $r$  with fewer function evaluations.
- In addition, we can analyze in more detail the set of solutions for  $\{a_i, b_i\}$ , and find very efficient methods

# Positive coefficients?

- When order  $r \geq 3$ , splitting methods (5) involve some negative coefficients
- They cannot be used for certain problems, e.g., ODEs defined in semigroups
- This result can be established as a theorem (Sheng, Suzuki). For an elementary proof, [Blanes-C.](#)
- Question: is this also true for processed methods?
- The theorem applies to the whole composition  $\hat{\psi}_h = \pi_h \psi_h \pi_h^{-1}$ , but may be  $\psi_h$  could involve only positive coefficients.
- The answer is **no**.

# Positive coefficients?

- When order  $r \geq 3$ , splitting methods (5) involve some negative coefficients
- They cannot be used for certain problems, e.g., ODEs defined in semigroups
- This result can be established as a theorem (Sheng, Suzuki). For an elementary proof, [Blanes-C.](#)
- Question: is this also true for processed methods?
- The theorem applies to the whole composition  $\hat{\psi}_h = \pi_h \psi_h \pi_h^{-1}$ , but may be  $\psi_h$  could involve only positive coefficients.
- The answer is **no**.

# Negative coefficients in the kernel

## Theorem

*At least one of the  $a_i$  as well as one of the  $b_i$  coefficients have to be negative in the composition*

$$\psi_h = e^{ha_1 F^{[a]}} e^{hb_1 F^{[b]}} e^{ha_2 F^{[a]}} \dots e^{ha_p F^{[a]}} e^{hb_p F^{[b]}}$$

*if  $\psi_h$  is the kernel of a processed method of order (or equivalently if  $\psi_h$  is of effective order)  $r \geq 3$ .*

(Blanes-C, Chin 2005)

- Thus the presence of negative coefficients in kernels of effective order  $r \geq 3$  is unavoidable.
- Is it possible to circumvent this (fundamental) difficulty?

## One (partial) remedy: 'modified potentials'

- For Hamiltonian systems  $H(p, q) = \frac{1}{2}p^T M p + V(q)$ , then  $F^{[a]} = J \nabla T(p)$  and  $F^{[b]} = J \nabla V(q)$  and

$$\begin{aligned} e^{haF^{[a]}}(x_0) &= (q_0 + h a M^{-1} p_0, p_0)^T \\ e^{hbF^{[b]}}(x_0) &= (q_0, p_0 - h b \nabla V(q_0))^T. \end{aligned} \quad (6)$$

- If  $[F^{[a]}, F^{[b]}] = F^{[a]}F^{[b]} - F^{[b]}F^{[a]}$ , then

$$F^{[bab]} \equiv [F^{[b]}, [F^{[a]}, F^{[b]}]] = -J(\nabla V)^T M^{-1} \nabla V$$

depends only on  $q$

- Then  $[F^{[b]}, [F^{[b]}, [F^{[a]}, F^{[b]}]] = 0$ , the flow of  $F^{[bab]}$  is explicitly computable and can be included in the integrator

## Examples of 'modified potentials'

- Alternatively, replace  $\exp(hb_j F^{[b]})$  by  $\exp(hF_{b_j, c_j}^{[b, c]})$ , where

$$F_{b_j, c_j}^{[b, c]} \equiv b_j F^{[b]} + c_j h^2 [F^{[b]}, [F^{[a]}, F^{[b]}]]$$

corresponding to the 'modified potential'

$$W_{b_j, c_j} = b_j V(q) - c_j h^2 (\nabla V(q))^T M^{-1} \nabla V(q).$$

- Order 4

$$\psi_h = e^{\frac{h}{6} F^{[b]}} e^{\frac{h}{2} F^{[a]}} e^{h F_{2/3, 1/72}^{[b, c]}} e^{\frac{h}{2} F^{[a]}} e^{\frac{h}{6} F^{[b]}}$$

(Koseleff, Chin)

- Effective order 4

$$\psi_h = e^{\frac{h}{2} F^{[a]}} e^{h F_{1, 1/24}^{[b, c]}} e^{\frac{h}{2} F^{[a]}}$$

(Takahashi-Imada, Rowland)

- If modified potentials of higher degree are considered, then it is possible to build methods of order 6 and higher with  $a_i$ ,  $b_i$  positive (Blanes-C.-Ros), e.g. the flow of

$$F^{[d]} \equiv [F^{[b]}, [F^{[b]}, [F^{[a]}, [F^{[a]}, F^{[b]}]]]]$$

- Unfortunately, for some problems the computation of these high-degree modified potentials is not feasible.

Open question: Is it possible to construct methods of effective order 6 involving only the modified potential

$W_{b_j, c_j} \equiv b_j V(q) + c_j h^2 (\nabla V(q))^T \nabla V(q)$  with all the  $a_i$  and  $b_i$  coefficients being positive?

- Currently, there are no known methods with these features.

# PIMC

- In condensed matter physics, splitting methods with positive coefficients could be extremely useful in **path integral Monte Carlo simulations**
- To study theoretically the thermal properties of a quantum system described by the Hamiltonian  $\mathcal{H}$  at finite temperature, one has to compute the **quantum partition function**

$$Z = \text{Tr}(e^{-\beta\mathcal{H}}) = \sum_{\alpha} \langle \alpha | e^{-\beta\mathcal{H}} | \alpha \rangle,$$

for the complete, orthonormal set  $|\alpha\rangle$ . Here  $\beta = k_B/T$

- This can be done with the Feynman path integral approach.

## PIMC

Within this approach,

- Partition function:  $Z = \int d\mathbf{R} \langle \mathbf{R} | e^{-\beta\mathcal{H}} | \mathbf{R} \rangle \equiv \int d\mathbf{R} \rho(\mathbf{R}, \mathbf{R}; \beta)$
- Density matrix:  $\rho(\mathbf{R}, \mathbf{R}'; \beta) \equiv \langle \mathbf{R} | e^{-\beta\mathcal{H}} | \mathbf{R}' \rangle$
- Procedure:
  - Factorize  $\exp(-\beta\mathcal{H}) = (\exp(-\varepsilon\mathcal{H}))^M$  with  $\varepsilon = \beta/M$  and a positive integer  $M$  (number of *beads*)
  - Approximate  $\exp(-\varepsilon\mathcal{H})$
- Then

$$Z = \int \cdots \int d\mathbf{R}_0 d\mathbf{R}_1 d\mathbf{R}_2 \cdots d\mathbf{R}_{M-1} \rho(\mathbf{R}_0, \mathbf{R}_1; \varepsilon) \times \rho(\mathbf{R}_1, \mathbf{R}_2; \varepsilon) \cdots \rho(\mathbf{R}_{M-1}, \mathbf{R}_0; \varepsilon), \quad (7)$$

# PIMC

- Typically,  $\mathcal{H} = \hat{K} + \hat{V}$ : kinetic + potential energy
- Density matrix corresponding to  $\hat{K}$  and  $\hat{V}$  can be computed explicitly
- Approximate  $e^{-\varepsilon\mathcal{H}}$  by symmetric products of  $e^{-\varepsilon\hat{K}}$  and  $e^{-\varepsilon\hat{V}}$
- Simplest approximation (**primitive action** PA):

$$e^{-\varepsilon(\hat{K}+\hat{V})} \simeq e^{-\varepsilon\hat{K}} e^{-\varepsilon\hat{V}}, \quad (8)$$

- First order in  $\varepsilon = \beta/M$  (effective order 2)
- To study a fully quantum regime at very low temperatures,  $M$  increases very fast and the efficiency of the simulation suffers.
- Higher actions (higher order approximations in  $\varepsilon$ ) are required.

## Higher order actions

- Symmetric factorizations of the form

$$e^{h(\hat{K}+\hat{V})} \simeq \prod_{i=1}^P e^{a_i h \hat{K}} e^{b_i h \hat{V}}, \quad (9)$$

where  $h \equiv -\varepsilon$  and coefficients  $\{a_i, b_i\}$  to be determined.

- All these parameters must be positive.
- Only in this way the integrand in  $Z$  can be normalized as a probability distribution and standard Metropolis Monte Carlo methods can be used
- $\implies$  we must introduce modified potentials  $\hat{W}_{b_i, c_i} = b_i \hat{V} + c_i h^2 [\hat{V}, [\hat{K}, \hat{V}]]$

Important feature: the trace in  $Z$  is invariant under similarity, so that we may construct only the kernel of a processed scheme to approximate  $e^{-\beta H}$  (*no processor is necessary*).

# Higher order actions

- The method  $\psi_h = e^{\frac{h}{2}F[a]} e^{hF_{1,1/24}^{[b,c]}} e^{\frac{h}{2}F[a]}$  leads to  $Z$  accurate up to fourth order (**Takahashi–Imada action**).
- Still more accuracy is necessary to deal properly with fully quantum fluids at ultra low temperature.
- Real fourth-order actions, possibly (hopefully!) of effective order 6.

It makes sense then to analyze whether it is possible to design methods of order four and effective order six containing only positive  $\{a_i, b_i\}$  coefficients in the composition.

# A particular method

- Recently, the 4th-order symmetric kernel ([Scuro-Chin'05](#))

$$\psi_h = e^{ha_1 \hat{K}} e^{h\hat{W}_{b_1, c_1}} e^{ha_2 \hat{K}} e^{h\hat{W}_{b_2, c_2}} e^{ha_2 \hat{K}} e^{h\hat{W}_{b_1, c_1}} e^{ha_1 \hat{K}}$$

has been analyzed in quantum MC simulations ([Sakkos-Casulleras-Boronat, 2009](#)).

- Two free parameters:  $0 \leq \alpha \leq \frac{1}{2}$ ,  $0 \leq t_0 \leq \frac{3-\sqrt{3}}{6}$

$$a_1 = t_0, \quad b_1 = \frac{1}{6(1-2t_0)^2}, \quad c_1 = u_0 \alpha$$

$$a_2 = \frac{1}{2} - t_0, \quad b_2 = 1 - 2b_1, \quad c_2 = u_0(1 - 2\alpha)$$

$$u_0 = \frac{1}{12} \left( 1 - \frac{1}{1-2t_0} + \frac{1}{6(1-2t_0)^3} \right)$$

so that all the  $a_i$ ,  $b_i$  are positive.

## Experimental conclusions

- The required  $M$  to reproduce the exact energy is much smaller than with PA and TIA.
- Empirical procedure for adjusting  $\alpha$  and  $t_0$  to improve accuracy from 4th-order to 6th-order
- Error minimized with  $\alpha = 0.33$  and  $t_0 = 0.1215$ . This is exact for the harmonic oscillator and remains approximately true for other systems ( $\text{H}_2$  drop and liquid  $^4\text{He}$ ).
- Goals:
  - Explain these observed phenomena
  - Construct new schemes with better efficiency on an enlarged range of values of  $\varepsilon$

## Analysis of a family of 4th-order actions

- Ansatz: symmetric composition

$$\psi_h = \prod_{i=1}^p \exp(a_i h \hat{K}) \exp(h \hat{W}_{b_i, c_i})$$

- Order conditions (up to  $r = 4$ ):

$$\sum_{i=1}^p a_i = 1, \quad \sum_{i=1}^p b_i = 1, \quad \sum_{i=1}^p b_i \left( \sum_{j=1}^i a_j \right)^2 = \frac{1}{3}$$

$$\frac{1}{2} \sum_{i=1}^p a_i \left( \sum_{j=i}^p b_j \right)^2 - \sum_{i=1}^p c_i = \frac{1}{6}$$

- Two more equations to achieve order 6:

$$G_1 \equiv 2\alpha_1 - \alpha_2 - \frac{1}{120} = 0, \quad G_2 \equiv \alpha_3 + \alpha_4 - \frac{1}{60} = 0$$

- $\alpha_1, \alpha_2$ : polynomials in terms of  $a_i, b_i$
- $\alpha_3, \alpha_4$  depending linearly on  $c_j$ .
- The analysis is simpler in terms of

$$s_i \equiv \sum_{j=1}^i a_j, \quad s_0 \equiv 0, \quad s_p = 1,$$

- We only analyze equations involving only  $a_i, b_i$ , i.e.,

$$\sum_{i=1}^p b_i = 1, \quad \sum_{i=1}^p s_i^2 b_i = \frac{1}{3}$$

$$\sum_{i=1}^p s_i^4 b_i - \frac{1}{3} \sum_{i=1}^{p-1} \sum_{j=i+1}^p (s_j - s_i)^3 b_i b_j = \frac{1}{60}$$

- Two types of symmetric compositions:
  - ABA-composition:  $b_p = 0$ ,  $b_{p-i} = b_i$ ,  $s_{p-i} = 1 - s_i$ .
  - BAB-composition:  $b_{p+1-i} = b_i$ ,  $s_{p-i} = 1 - s_{i+1}$
- $c_i$  coefficients are related in a similar way as the  $b_i$
- $a_i$  are non-negative as long as  $s_0 = 0 \leq s_1 \leq s_2 \leq \dots \leq s_p = 1$

### ABA-compositions

- $p = 4$  is the minimum number. This corresponds to

$$\psi_h = e^{ha_1K} e^{hW_{b_1,c_1}} e^{ha_2K} e^{hW_{b_2,c_2}} e^{ha_2K} e^{hW_{b_1,c_1}} e^{ha_1K}$$

- Three solutions for  $b_1$ ,  $b_2$  and  $s_1$ , but only one real:  
 $s_1 \simeq 0.486421$ . All  $a_i$  coefficients positive, but  $b_1 = 225.975$ ,  
 $b_2 = -450.949$ .
- At least one  $b_j$  negative.

# ABA $p = 4$

- When  $\hat{V} = \frac{1}{2}\lambda^2 x^2$ , it is possible to get ABA methods with  $p = 4$  of order 4 and effective order 6: if

$$t_0 = s_1 \in (0, 0.132024) \cup (0.142326, 0.180679)$$

then all the coefficients are positive and the schemes are of effective order 6 for the harmonic oscillator.

- For  $s_1 = t_0 = 0.1215$ , one gets  $\alpha = 0.329556$ .
- This explains the observed behavior
- But the composition is no longer of effective order 6 for a general potential.
- Same conclusions are derived for  $p = 5$  and  $p = 6$ .

# BAB methods

- Here also  $p = 4$
- Family of methods:

$$\psi_h = e^{h\hat{W}_{b_1, c_1}} e^{ha_1\hat{K}} e^{h\hat{W}_{b_2, c_2}} e^{ha_2\hat{K}} e^{h\hat{W}_{b_2, c_2}} e^{ha_1\hat{K}} e^{h\hat{W}_{b_1, c_1}}. \quad (10)$$

- FSAL property: 3 stages (same as before)
- 3 variables to analyze:  $b_1, b_2, a_1 = s_2$
- All coefficients are positive when  $s_2 \in (0.211325, 0.350226) \cup (0.459992, 0.5)$  for the harmonic oscillator
- No longer true for a general potential
- Same behavior for  $p \geq 5$ .

# Linear stability

- In addition to order of accuracy, **stability**
- Stable numerical integrator: the numerical solution does not tend to infinity when the exact solution is bounded
- Numerical integrators with a large stability interval are suitable in PIMC: in this way  $h$  is as large as possible,  $M$  is small and the computational complexity is reduced
- Linear stability is analyzed on the 1-dimensional harmonic oscillator

$$y'' + \lambda^2 y = 0, \quad \lambda > 0,$$

$$\text{with } x = (q, p)^T = (q, p) = (\lambda y, y')^T$$

$$\begin{pmatrix} q' \\ p' \end{pmatrix} = \left[ \underbrace{\begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix}}_A + \underbrace{\begin{pmatrix} 0 & 0 \\ -\lambda & 0 \end{pmatrix}}_B \right] \begin{pmatrix} q \\ p \end{pmatrix},$$

# Linear stability

- Exact solution:

$$O(x) = \begin{pmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{pmatrix}, \quad x = h\lambda,$$

- Approximation:

$$\begin{aligned} K(x) &= \begin{pmatrix} 1 & a_1 x \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -x(b_1 - 2c_1 x^2) & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & a_p x \\ 0 & 1 \end{pmatrix} \\ &\quad \times \begin{pmatrix} 1 & 0 \\ -x(b_p - 2c_p x^2) & 1 \end{pmatrix} \\ &= e^{ha_1 \hat{K}} e^{h\hat{W}_{b_1, c_1}} \cdots e^{ha_p \hat{K}} e^{h\hat{W}_{b_p, c_p}}. \end{aligned}$$

- Goal: find the maximal value of  $x$  for which  $K(x)^n$  remain bounded for all  $n$

- Stability threshold: the largest non-negative real number such that  $K(x)$  is stable for all  $x \in (-x_*, x_*)$
- $K(x)^n$  can be bounded independently of  $n \geq 1$  for  $x \in (-x_*, x_*)$  if all the eigenvalues of  $K$  lie on the stability interval  $(-x_*, x_*)$ .
- $x_* \leq 2p$  since  $2p$  is the maximal value of the stability threshold, which is achieved by the concatenation of  $p$  steps of length  $h/p$  of the leapfrog scheme
- For processed methods the linear stability is determined only by the kernel

- An essential role is played by the **stability polynomial**  
$$p(x) = \frac{1}{2} \text{Tr } K(x)$$
- If the method is stable for a given  $x \in \mathbb{R}$ , then  $|p(x)| \leq 1$
- Practical criterion for estimating  $x_*$  based on computing the positive zeros of  $p(x)^2 - 1$
- We analyze the linear stability of the 1-parameter 4th-order actions considered before
- Idea: to determine for each family the stability threshold  $x_*$  as a function of the free parameter and select the parameter leading to the largest value of  $x_*$ .

- For both families

$$p(x) = 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 - \frac{1}{720}x^6 + \mathcal{O}(x^8),$$

- Dependence on  $s_1$  ( $s_2$ ) at higher order terms
- $p(x)$  approximates  $\cos x$  up to order 6, in accordance with the order of accuracy of the schemes
- Results:

Method	Parameter	Largest $x_*$	$x_*/p$
ABA	$s_1 = 0.1234$	3.0731	1.02436
BAB	$s_2 = 0.2785$	3.1399	1.04663

- $x_*/p$  should be compared with 2, the maximal stability threshold attained by leapfrog

# Remarks

- With the BAB composition it is possible to achieve a larger stability interval: **new method which can be used with a larger step size**
- For the value considered by Boronat *et al.* for ABA,  $s_1 = t_0 = 0.1215$ , one has  $x_*/p = 1.02419$
- It is possible to adjust the parameter to get higher stability

## Conclusions:

- We have provided a sound justification of previous empirical observations
- We have designed new methods within this family with the same computational cost and better properties (larger time steps)
- It is not strictly necessary to satisfy exactly all the effective order conditions at order 6, only in the linear case
- Important: to have a large stability threshold
- Essential: positive coefficients

## Outlook:

- Design new methods with more stages than strictly necessary (e.g.,  $p = 5$  or  $p = 6$ ) and use the extra parameters to fulfill effective order conditions at higher orders for the linear case, whereas
  - keeping all the coefficients positive
  - ensuring a large stability interval
- Apply the new methods in actual Monte Carlo simulations
- Consider other settings: Schrödinger eq. with imaginary time, and diffusion equations.