WORD SERIES

FOR THE ANALYSIS OF SPLITTING SDE INTEGRATORS

Alfonso Álamo/J. M. Sanz-Serna

Universidad de Valladolid/Universidad Carlos III de Madrid

I. OVERVIEW

- The importance of splitting integrators for

$$(d/dt)x = f_a(x) + f_b(x) + \cdots$$

keeps increasing: evolutionary PDEs, geometric integration, ...

- Format: maps that advances one step given by composition

$$\psi_h = \phi^a_{\lambda h} \circ \phi^b_{\mu h} \circ \cdots ,$$

where h is the stepsize and ϕ^a_h, ϕ^b_h, ... are exact solution flows of $(d/dt)x = f_a(x)$, $(d/dt)x = f_b(x)$, ...

• Standard analysis of ψ_h starts by seeing ϕ_h^a, ϕ_h^b, ... as (non-commuting) exponentials.

• BCH formula then used to write ψ_h as a single exponential. The exponent provides the modified differential equation of the integrator, ie the (h-dependent) equation whose h flow formally coincides with ψ_h.

• Difference between modified and true $(d/dt)x = f_a(x) + f_b(x) + \cdots$ equations yields information on the properties of the integrator.

• This is an indirect approach based on comparing differential equations rather than maps.

• It stands apart from usual techniques in numerical ODEs based on expansion of integrator map ψ_h.

• For Runge-Kutta and related integrators the expansion of ψ_h is best carried out by means of B-series (Hairer and Wanner 1974). B-series are parameterized by rooted trees.

• Murua and SS (1999) developed a B-series technique to analyze splitting integrators (Hairer, Lubich, Wanner, GI Book, Chapter III.3).

• Recently Murua and SS suggested word series as an alternative to the B-series. Word series are parameterized by words from an alphabet.

• Use of word series based on simple algebraic systematic manipulations.

- In numerical analysis, word series may be used to study the order of consistency of splitting integrators, find modified equations, etc.

- Outside numerical mathematics, word series may be used (Chartier, Murua, SS 2010–2015) to perform averaging, to find normal forms, to compute integrals of motion, etc.

- In this talk we present an example of the use of word series to analyze splitting methods for SDEs. (Also: Ito SDEs, modified equations, . . .)

II. DEFINING WORD SERIES

A MOTIVATING (DETERMINISTIC) EXAMPLE:

$$dx = \sum_{a \in \mathcal{A}} \lambda_a(t)\, dt\, f_a(x).$$

- $\mathcal{A}$ finite or infinite index set (alphabet).
- $f_a(x)$ vector field in D dimensions.
- $\lambda_a(t)$ scalar-valued function.

- Solution with $x(0) = x_0$ given by (cf Chen series)

$$x(t) = x_0 + \sum_{n=1}^{\infty} \sum_{a_1,\ldots,a_n \in A} J_{a_1\cdots a_n}(t) f_{a_1\cdots a_n}(x_0),$$

with $J_{a_1\cdots a_n}(t)$ given by

$$\int_0^t \lambda_{a_n}(t_n)\, dt_n \int_0^{t_n} \lambda_{a_{n-1}}(t_{n-1})\, dt_{n-1} \cdots \int_0^{t_2} \lambda_{a_1}(t_1)\, dt_1,$$

and, recursively,

$$f_{a_1\cdots a_n}(x) = \partial_x f_{a_2\cdots a_n}(x)\, f_{a_1}(x).$$

- Note separation: λ_a, $J_{a_1\cdots a_n}$ / f_a, $f_{a_1\cdots a_n}$.

WORDS AND THEIR BASIS FUNCTIONS

• $\mathcal{W}$ denotes set of all words from the alphabet $\mathcal{A}$ (including the empty word $\emptyset$); $\mathcal{W}_n$ is the set of words with n letters.

• Associate with $w = a_1 \cdots a_n \in \mathcal{W}$ its word basis function $f_w(x) = f_{a_1 \cdots a_n}(x)$ (with $f_\emptyset(x) \equiv x$). (Similar to elementary differentials.)

DEFINITION

• If $\delta \in \mathbb{C}^{\mathcal{W}}$ (ie δ maps words into scalars), the word series with coefficients δ is the formal series:

$$W_\delta(x) = \sum_{w\in\mathcal{W}} \delta_w f_w(x).$$

• Note that the notion of word series is relative to the f_a.

• For each t, the solution value $x(t)$ corresponds to the coefficients $J_w(t)$ built above from the $\lambda_a(t)$.

SOME PARTICULAR CASES:

EXAMPLE I: $\mathcal{A}$ consists of a single letter a. For each n there is a single word $w = a \cdots a$ with n letters.

If furthermore $\lambda_a(t) \equiv 1$, $J_w(t)$, $w \in \mathcal{W}_n$ is found to be $t^n/n!$. The word series $W_{\alpha(t)}(x_0)$ is the Taylor expansion in powers of t of the solution $x(t)$ of $(d/dt)x = f_a(x)$, $x(0) = x_0$.

EXAMPLE II: $\mathcal{A}$ consists of a two letters a, b. For each n there are 2^n words with n letters.

If furthermore $\lambda_a(t) \equiv 1$, $\lambda_b(t) \equiv 1$, $\alpha_w(t) = t^n/n!$ for $w \in \mathcal{W}_n$. The word series $W_{J(t)}(x_0)$ is the Taylor expansion of $x(t)$, $(d/dt)x = f_a(x) + f_b(x)$, written in terms of f_a, f_b.

STOCHASTIC SDEs:

- $dx = f_a(x)dt + f_A dB(t)$ (Stratonovich), fits in this framework with $\lambda_a(t) \equiv 1$ and $\lambda_A(t)dt = dB(t)$. Solution formally given by $W_{J(t)}(x_0)$ where now $J_w(t)$ is a stochastic process. (Stochastic Taylor expansion see Kloeden and Platen, Chapter 5.)

III. OPERATING WITH WORD SERIES

THE CONVOLUTION PRODUCT $\star$

- If $\delta, \delta' \in \mathbb{C}^{\mathcal{W}}$, their *convolution product* $\delta \star \delta' \in \mathbb{C}^{\mathcal{W}}$ is

$$(\delta \star \delta')_{a_1 \cdots a_n} = \delta_{\emptyset} \delta'_{a_1 \cdots a_n} + \sum_{j=1}^{n-1} \delta_{a_1 \cdots a_j} \delta'_{a_{j+1} \cdots a_n} + \delta_{a_1 \cdots a_n} \delta'_{\emptyset}$$

- Not commutative.

- Associative.

- Unit: $\mathbb{1} \in \mathbb{C}^{\mathcal{W}}$ with $\mathbb{1}_{\emptyset} = 1$ and $\mathbb{1}_{w} = 0$ for $w \neq \emptyset$.

THE GROUP $\mathcal{G}$ (Group of characters of shuffle Hopf algebra)

• Let $\sqcup\!\sqcup$ denote shuffle product of words. (Eg $ab \sqcup\!\sqcup c = abc + acb + cab$.)

• The set $\mathcal{G}$ of those $\gamma \in \mathbb{C}^{\mathcal{W}}$ that satisfy the so-called *shuffle relations:* $\gamma_\emptyset = 1$ and, for each $w, w' \in \mathcal{W}$,

$$\gamma_w \gamma_{w'} = \sum_{j=1}^{N} \gamma_{w_j} \qquad \text{if} \qquad w \sqcup\!\sqcup w' = \sum_{j=1}^{N} w_j.$$

is a noncommutative group for $\star$.

• For fixed t (and $\omega \in \Omega$), the coefficients $J_w(t)$, $w \in \mathcal{W}$ satisfy the shuffle relations, ie they give an element of $\mathcal{G}$.

• For $\gamma \in \mathcal{G}$, the series $W_\gamma(x)$ has special properties:

(1) $W_\gamma(x)$ acts on the vector space of all word series by composition:

$$W_\delta\big(W_\gamma(x)\big) = W_{\gamma\star\delta}(x), \quad \delta \in \mathbb{C}^{\mathcal{W}}.$$

(2) If $\chi : \mathbb{R}^D \to \mathbb{R}$ is any smooth observable:

$$\chi\big(W_\gamma(x)\big) = \sum_{w \in \mathcal{W}} \gamma_w D_w(\chi)(x),$$

where, for $a \in \mathcal{A}$, $D_a = f_a \cdot \nabla$ is the Lie operator associated with f_a and for $w = a_1 \ldots a_n$, D_w is the differential operator $D_{a_1} \cdot \ldots \cdot D_{a_n}$.

(3) Equivariance with respect to arbitrary changes of variables $x = x(X)$.

IIII. SPLITTING INTEGRATORS FOR SDES

• Consider integrators for Stratonovich equation

$$\begin{aligned} dx &= f_{a_1}(x)dt + \cdots + f_{a_n}(x)dt \\ &\quad + f_{A_1}(x)dB_{A_1} + \cdots + f_{A_N}(x)dB_{A_N} \end{aligned}$$

such that the one-step map ψ_h is composition of flows of individual pieces $dx = f_{a_i}(x)dt$ or $dx = f_{A_j}(x)dB_{A_j}$ (or blocks, eg $dx = f_{a_1}(x)dt + f_{A_2}(x)dB_{A_j}$).

• Each flow entering the composition is (for fixed t and $\omega \in \Omega$) a word series with known coefficients in $\mathcal{G}$.

- By property (1), $\psi_h = W_{\gamma(t)}(x)$, where the coefficients $\gamma_w(t)$ are readily found by using the convolution formula.

- $\gamma_w(t)$ compared with coefficients $J_w(t)$ of true solution. For strong local error to be $\mathcal{O}(h^p)$, $p = 3/2, 2, 5/2, \ldots$

$$\gamma_w(t) = J_w(t) \qquad \text{as}$$

for all words of weight $\leq p - 1/2$ (weight = number of lower case letters + number of upper case letters/2). (This assumes all relevant basis functions are $\neq 0$.)

- By property (3)

$$\mathbb{E}\Big(\chi\big(W_\gamma(t)(x)\big)\Big) = \sum_{w\in\mathcal{W}} \mathbb{E}\big(\gamma_w(t)\big) D_w(\chi)(x),$$

- For weak local error to be $\mathcal{O}(h^p)$, $p = 2, 3, \ldots$,

$$\mathbb{E}\big(\gamma_w(t)\big) = \mathbb{E}\big(J_w(t)\big),$$

for all words of weight $\leq p - 1$. (This assumes all relevant basis functions are $\neq 0$.)

V. LANGEVIN DYNAMICS

- Consider

$$\begin{aligned} dq &= M^{-1} p \, dt \\ dp &= F(q) \, dt - \gamma p \, dt + \sigma M^{1/2} dB, \end{aligned}$$

M diagonal with diagonal elements $m_i > 0$, $i = 1, \ldots, d$, $\gamma > 0$, B d-dimensional Wiener process.

• Leimkuhler and Matthews (2013) consider several integrators based on the split systems:

a: $dq = M^{-1}p\,dt$, $dp = 0$.

b: $dq = 0$, $dp = F(q)\,dt$.

o: $dq = 0$, $dp = -\gamma p\,dt + \sigma M^{1/2} dB$.

• They use acronym aboba for the Strang-like splitting

$$\psi_h = \phi^a_{h/2} \circ \phi^b_{h/2} \circ \phi^o_h \circ \phi^b_{h/2} \circ \phi^a_{h/2};$$

baoab defined similarly.

• Both aboba and baoab possess strong local error $\mathcal{O}(h^{3/2})$; weak $\mathcal{O}(h^3)$. In spite of similarity, baoab turns out to be clearly superior to aboba.

- We use word series based on the vector fields:

a: $(M^{-1}p, 0)$.

b: $(0, F(q))$.

c: $(0, -\gamma p)$.

A_j: $(0, \sigma\sqrt{m_j}e_j), \quad j = 1, \dots, d$.

• Sparsity pattern results in many zero word basis functions.

• For weight < 3, $f_w \neq 0$ only for A_j, A_ja, A_jc, A_jab, A_jca, A_jcc and some purely deterministic words (that cause no trouble).

• The coefficients of the methods in the word series expansion are easily found to be:

w	weight	aboba	baoab
A_j	1/2	$\checkmark$	$\checkmark$
A_ja	3/2	$(h/2)J_{A_j}$	$(h/2)J_{A_j}$
A_jc	3/2	$\checkmark$	$\checkmark$
A_jab	5/2	0	$(h^2/4)J_{A_j}$
A_jca	5/2	$(h/2)J_{A_jc}$	$(h/2)J_{A_jc}$
A_jcc	5/2	$\checkmark$	$\checkmark$

- Only difference in A_jab, where baoab provides a better approximation: correlation truth-approximation is $\sqrt{5}/3 \approx 74\%$ for baoab, 0 for aboba.

- Symmetrically, for word $A_j ba$, baoab has $\gamma_{A_j ba} = 0$, while aboba has $\gamma_{A_j ba} = (h^2/4) J_{A_j}$. But then the basis function is zero and aboba does not benefit from it.

- The same phenomena appear for longer words.

- Word analysis reveals that the source of discrepancy is the following algorithmic flaw in aboba: in any given time step, it uses the *same* value of the force in both kicks of a given step.