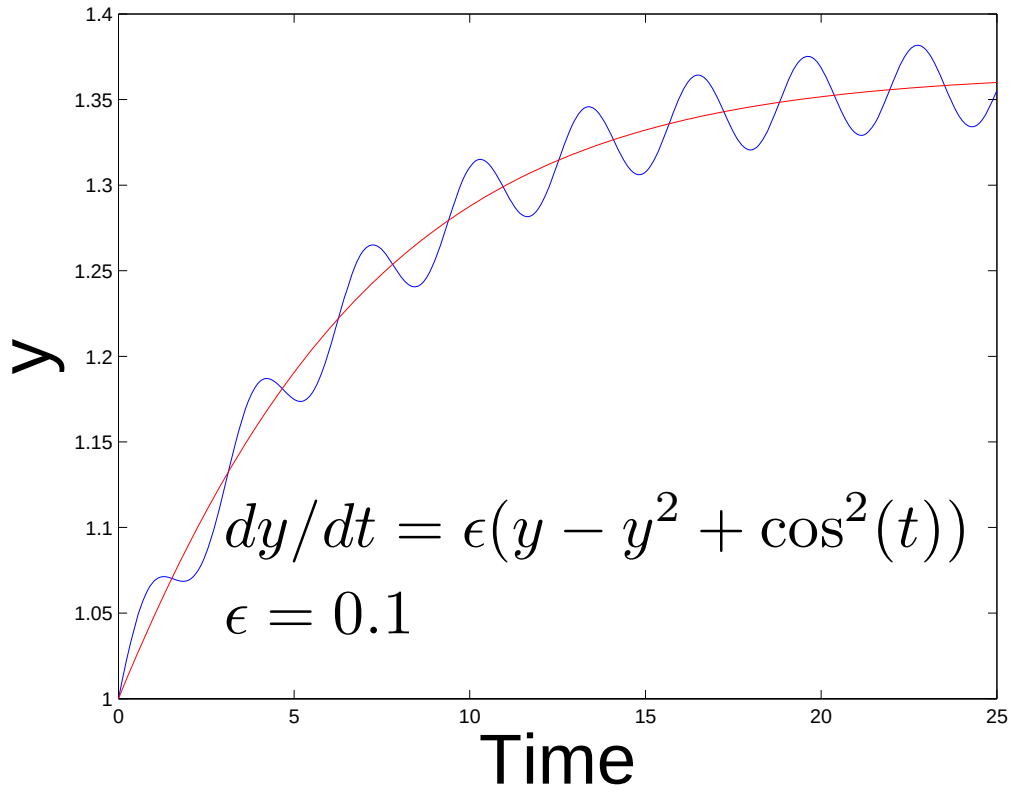
Averaging old and new

J. M. Sanz-Serna
Universidad de Valladolid, Spain

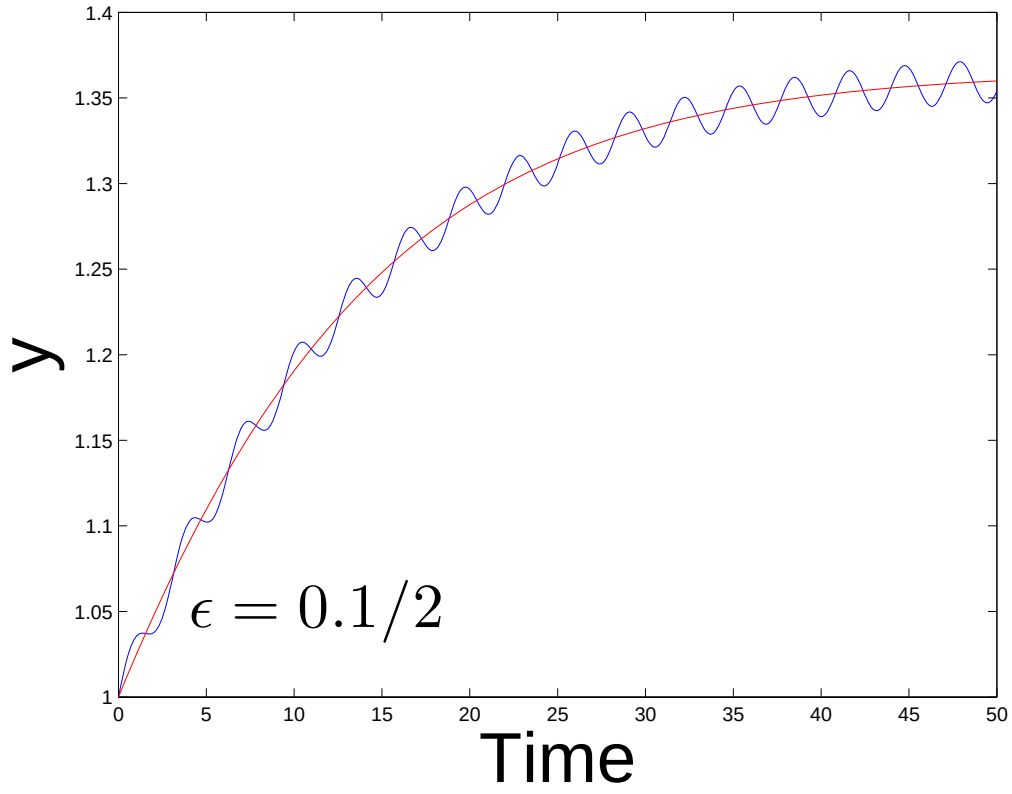
Joint work with A. Murua & P. Chartier

I. AVERAGING

- An easy example (in red *averaged solution* where $\cos^2(t) = 1/2(1 + \cos(2t))$ has been replaced by its average $1/2$) ($0 \leq t \leq 25$):



- Divide ϵ by 2 and integrate on twice-as-long time-interval ($0 \leq t \leq 50$):



- *Averaging yields $\mathcal{O}(\epsilon)$ errors over t-intervals of length $\mathcal{O}(1/\epsilon)$.*

• First used in eighteenth century (e.g. Lagrange, Laplace, . . .)

• Often (and not always correctly) used since then in celestial mechanics and many other application fields.

• Krylov & Bogoliubov 1937, Bogoliubov and & Mitropolskii 1961.

• Many variants and extensions, e.g. right-hand side (vector field) $f(y,t)$ not periodic in t.

• High-order versions, where true solution $y(t)$ is approximated to order $\mathcal{O}(\epsilon^N)$, $N > 1$, over t-intervals of length $\mathcal{O}(1/\epsilon)$.

• Consider hereafter the initial value problem ($0 < \varepsilon \ll 1$):

$$\frac{d}{dt}y = \varepsilon f(y, t\omega), \qquad 0 \le t \le L/\varepsilon, \qquad y(0) = y_0 \in \mathbb{R}^D.$$

• $f(y, \theta)$ 2π-periodic wrt each scalar components θ_1,...,θ_d of the angular variable $\theta \in \mathbb{T}^d$:

$$f(y, \theta) = \sum_{\mathbf{k} \in \mathbb{Z}^d} e^{i\mathbf{k}\cdot\theta} f_{\mathbf{k}}(y).$$

If $d = 1$, $f(y, t\omega)$ is $2\pi/\omega$ *periodic* in t.

If $d > 1$, problem is *quasi-periodic*. Assume then ω non-resonant: $\mathbf{k}\cdot\omega \neq 0$ for each multi-index $\mathbf{k} \in \mathbb{Z}^d$, with $\mathbf{k} \neq \mathbf{0}$.

• Many problems may be brought to format above via suitable changes of variables. (More on this later.)

Theorem (Perko 1968) (Some technical assumptions required.) Given $N = 1, 2, \ldots$ solution $y(t)$ of ivp may be approximated with $\mathcal{O}(\epsilon^N)$, errors over t-intervals of length $\mathcal{O}(1/\epsilon)$ by $U_N(Y_N(t), t\omega)$, where:

- U_N is a change of variables of the form

$$U_N(Y, \theta) = Y + \varepsilon u_1(Y, \theta) + \cdots + \varepsilon^{N-1} u_{N-1}(Y, \theta), \qquad \theta \in \mathbb{T}^d,$$

- $Y_N(t)$ is the solution of an *autonomous* problem (the averaged problem)

$$\frac{d}{dt} Y = \varepsilon F_1(Y) + \varepsilon^2 F_2(Y) + \cdots + \varepsilon^N F_N(Y), \qquad Y(0) = \xi,$$

(ξ is determined implicitly through the equation $U_N(\xi, 0) = y_0$).

• Functions F_1, u_1, F_2, u_2, F_3, ... determined recursively as follows.

• When F_1, u_1,..., F_{j-1}, u_{j-1}, have been found, the change of variables $Y + \varepsilon u_1(Y, t\omega) + \cdots + \varepsilon^j u_j(Y, t\omega)$ transforms the given system into

$$\frac{d}{dt}Y = \varepsilon F_1(Y) + \cdots + \varepsilon^i \Big(\tilde{F}_j(Y, t\omega) - \omega \cdot \nabla_\theta u_j(Y, t\omega) \Big) + \mathcal{O}(\epsilon^{i+1}),$$

where,

$$\tilde{F}_j(Y, \theta) = \sum_{r=1}^{j-1} \Big[\frac{1}{r!} \sum_{i_1 + \cdots + i_r = j-1} \frac{\partial^r f}{\partial y^r}(u_{i_1}, \ldots, u_{i_r}) - \frac{\partial u_r}{\partial Y} F_{j-r} \Big].$$

(For $j = 1$, $\tilde{F}_1(Y, \theta) := f(Y, \theta)$.)

• Since $u_j(Y, \theta)$ has to be 2π-periodic in the components of θ, no choice of u_j completely annihilates the red term.

• Choose

$$\begin{aligned} F_j(Y) &:= \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} \tilde{F}_j(Y,\theta)d\theta, \\ \omega\cdot\nabla_\theta u_j(Y,\theta) &= \tilde{F}_j(Y,\theta) - F_j(Y), \quad \text{(Homological equation)} \\ \int_{\mathbb{T}^d} u_j(Y,\theta)d\theta &= 0, \end{aligned}$$

so as to make the red term equal to $F_j(Y)$.

• Actual computation of the averaged system may be a daunting task (symbolic manipulation required, . . .).

• N may be chosen to be arbitrarily high; the asymptotic series

$$\varepsilon F_1(Y) + \varepsilon^2 F_2(Y) + \cdots ,$$

has the property that by keeping its first N terms the procedure yields an approximation to y with $\mathcal{O}(\epsilon^N)$ error in intervals of length $\mathcal{O}(1/\epsilon)$.

• However the series in general is not convergent: $y(t)$ cannot be *exactly* represented in terms of solutions $Y(t)$ of autonomous averaged problems.

• By choosing $N = N(\epsilon)$ cleverly one may approximate y up to *exponentially small errors* $\mathcal{O}(\exp(-k_1/\epsilon^{k_2}))$, $k_i > 0$, Neishtadt (1984) ($d = 1$), Simó (1994) ($d > 1$).

• Neishtadt's result crucial in symplectic integration.

II. A NEW APPROACH

(II.A) New approach begins by finding the formal expansion of $y(t)$ in powers of ϵ:

- Step 1: Rewrite as integral equation:

$$\begin{aligned} y(t) &= y_0 + \varepsilon \int_0^t f(y(s), s\omega)\, ds \\ &= y_0 + \varepsilon \int_0^t \sum_{\mathbf{k}\in\mathbb{Z}^d} e^{is\mathbf{k}\cdot\omega} f_{\mathbf{k}}(y(s))\, ds. \end{aligned}$$

- Step 2: Iterate *à la Picard*: $y(t) = y_0 + \mathcal{O}(\varepsilon) \Rightarrow$

$$y(t) = y_0 + \varepsilon \sum_{\mathbf{k}\in\mathbb{Z}^d} \left(\int_0^t e^{is\mathbf{k}\cdot\omega}\, ds \right) f_{\mathbf{k}}(y_0) + \mathcal{O}(\varepsilon^2),$$

$$\Rightarrow\ y(t) \;=\; y_0 + \varepsilon \sum_{\mathbf{k}\in\mathbb{Z}^d} \left(\int_0^t e^{is\mathbf{k}\cdot\omega}\, ds \right) f_{\mathbf{k}}(y_0)$$
$$+ \varepsilon^2 \sum_{\mathbf{k},\mathbf{l}\in\mathbb{Z}^d} \left(\int_0^t \int_0^{s_1} e^{i(s_1\mathbf{k}+s_2\mathbf{l})\cdot\omega}\, ds_1 ds_2 \right) f'_{\mathbf{k}}(y_0) f_{\mathbf{l}}(y_0)$$
$$+ \mathcal{O}(\varepsilon^3),$$

etc.

Note problem-dependent *elementary differentials*: $f_{\mathbf{k}}(y_0)$, $f'_{\mathbf{k}}(y_0)f_{\mathbf{l}}(y_0)$, ... and universal coefficients (independent of f, D). Let us systematize using some tools from numerical analysis.

(II.B) An aside: Analyzing numerical integrators:

• Investigation of accuracy of a linear multistep method (and hence construction of a high-order method) is trivial.

• For Runge-Kutta methods such a task, if performed naively (as in Runge or Kutta circa 1900), is overwhelming. Algebraic difficulties cloud the analysis.

• Combinatorial techniques developed in the 1960's (J. Butcher) to simplify and systematize study of integrators. These developments have been rediscovered and used later in other branches of maths (Hopf algebras, renormalization in quantum field theory, noncommutative geometry, . . .)

• Another milestone in the systematization: B-series, E. Hairer & G. Wanner 1974.

• B-series have been a major tool in Geometric Integration (SS 1996), starting with Calvo & SS 1994, Hairer 1994.

• Let us apply these tools to find the expansion of $y(t)$ in powers of ϵ.

(II.C) First tool: (mode-coloured rooted) trees

The *formal* definition of the set of trees $\mathcal{T}$ is recursive:

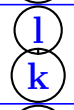
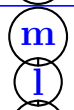
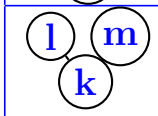
1. For each $\mathbf{k} \in \mathbb{Z}^d$, the tree $\bullet_{\mathbf{k}}$ (with one vertex —the root— coloured by $\mathbf{k}$) belongs to $\mathcal{T}$.

2. If $u_1, \ldots, u_n$ are trees $\in \mathcal{T}$, then the tree

$$u = [u_1 \cdots u_n]_{\mathbf{k}} \,,$$

obtained by connecting through an edge their roots to a new root labelled with the multi-index $\mathbf{k} \in \mathbb{Z}^d$, belongs to $\mathcal{T}$.

With ≤ 3 vertices:

u	$\mathcal{F}_u(y)$	$\alpha_u(t)$
	$f_{\mathbf{k}}(y)$	$\int_0^t e^{is_1\mathbf{k}\cdot\omega}\,ds_1$
	$f'_{\mathbf{k}}(y)f_{\mathbf{l}}(y)$	$\int_0^t\int_0^{s_1} e^{i(s_1\mathbf{k}+s_2\mathbf{l})\cdot\omega}\,ds_1\,ds_2$
	$f'_{\mathbf{k}}(y)f'_{\mathbf{l}}(y)f_{\mathbf{m}}(y)$	$\int_0^t\int_0^{s_1}\int_0^{s_2} e^{i(s_1\mathbf{k}+s_2\mathbf{l}+s_3\mathbf{m})\cdot\omega}\,ds_1\,ds_2\,ds_3$
	$f''_{\mathbf{k}}(y)(f_{\mathbf{l}}(y),f_{\mathbf{m}}(y))$	$\int_0^t e^{is_1\mathbf{k}\cdot\omega}\Big(\int_0^{s_1} e^{is_2\mathbf{l}\cdot\omega}\,ds_2\int_0^{s_1} e^{is_3\mathbf{m}\cdot\omega}\,ds_3\Big)ds_1$

With each tree $u \in \mathcal{T}$ we associate its *elementary differential* (relative to the functions $f_{\mathbf{k}}$): a map $\mathcal{F}_u : \mathbb{R}^D \to \mathbb{R}^D$ defined by:

1. For each tree $\textcircled{\mathbf{k}}$ of order 1, $\mathcal{F}_u(y) := f_{\mathbf{k}}(y)$.

2. For $u = [u_1 \cdots u_n]_{\mathbf{k}}$:

$$\mathcal{F}_u(y) := \frac{\partial^n f_{\mathbf{k}}}{\partial y^n}(y)\Big(\mathcal{F}_{u_1}(y), \ldots, \mathcal{F}_{u_n}(y)\Big).$$

With trees/elementary differentials as tools, we may find the expansion of $y(t)$ (second time):

• Use *ansatz*

$$y(t) = y_0 + \sum_{r=1}^{\infty} \varepsilon^r \sum_{|u|=r} \frac{1}{\sigma_u} \alpha_u(t) \mathcal{F}_u(y_0) = y_0 + \sum_{u \in \mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \alpha_u(t) \mathcal{F}_u(y_0),$$

where $\alpha_u(t)$ are coefficients to be determined and σ_u is a normalization factor (the number of symmetries of u).

• Take ansatz to integral formulation to find the following recursion to compute the $\alpha_u(t)$ . . .

For trees of order one, i.e. with one vertex:

$$\forall \mathbf{k} \in \mathbb{Z}^d, \qquad \alpha_{\textcircled{\mathbf{k}}}(t) = \int_0^t e^{is\mathbf{k}\cdot\omega}\, ds,$$

. . . and for larger trees . . .

$$\forall u = [u_1 \cdots u_n]_{\mathbf{k}}, \qquad \alpha_u(t) = \int_0^t e^{is\mathbf{k}\cdot\omega} \alpha_{u_1}(s) \cdots \alpha_{u_n}(s)\, ds.$$

(II.D) Second tool: (mode-coloured) B-series

• Analogy: each sequence of coefficients $\delta_1, \delta_2, \ldots$ determines a power series $\sum_n \delta_n z^n$. Operations for series (i.e. differentiation, composition,...) easily translated into operations for the sequences of coefficients.

• Similarly, with $\delta : \mathcal{T} \cup \{\emptyset\} \to \mathbb{C}$ we associate a series (B-series):

$$B(\delta, y) := \delta_\emptyset y + \sum_{r=1}^{\infty} \varepsilon^r \sum_{|u|=r} \frac{1}{\sigma_u} \delta_u \mathcal{F}_u(y) = \delta_\emptyset y + \sum_{u \in \mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \delta_u \mathcal{F}_u(y).$$

• Given B-series $B(\eta, y)$, $B(\delta, y)$, with $\delta_\emptyset = 1$, the composition

$$B(\eta, B(\delta, y))$$

is a new B-series $B(\zeta, y)$, where ζ *only depends on* δ *and* η (each ζ_u is a known polynomial in the δ_v's and η_w's). Write: $\zeta = \delta * \eta$.

Example:

$$\zeta\left(\begin{array}{ccc} \mathrm{l} & & \mathrm{m} \\ & \mathrm{k} & \end{array}\right) = \eta(\emptyset)\,\delta\left(\begin{array}{ccc} \mathrm{l} & & \mathrm{m} \\ & \mathrm{k} & \end{array}\right) + \eta\left(\begin{array}{c} \mathrm{k} \end{array}\right)\delta\left(\begin{array}{c} \mathrm{l} \end{array}\right)\delta\left(\begin{array}{c} \mathrm{m} \end{array}\right)$$
$$+ \eta\left(\begin{array}{c} \mathrm{l} \\ \mathrm{k} \end{array}\right)\delta\left(\begin{array}{c} \mathrm{m} \end{array}\right) + \eta\left(\begin{array}{c} \mathrm{m} \\ \mathrm{k} \end{array}\right)\delta\left(\begin{array}{c} \mathrm{l} \end{array}\right) + \eta\left(\begin{array}{ccc} \mathrm{l} & & \mathrm{m} \\ & \mathrm{k} & \end{array}\right)\delta(\emptyset).$$

(Note argument of η in rhs ranges over all trees obtaining by 'pruning' $\begin{array}{ccc} \mathrm{l} & & \mathrm{m} \\ & \mathrm{k} & \end{array}$.)

• The subset $\mathcal{G} \subset \mathbb{C}^{\mathcal{T}\cup\{\emptyset\}}$ of the mappings with $\delta_\emptyset = 1$ is a non-commutative group for the product $*$: the Butcher group. Its unit $\mathbb{1}$ has $\mathbb{1}_\emptyset = 1$, $\mathbb{1}_u = 0$.

With $\mathcal{G}$ as a tool, we may find the expansion of $y(t)$ (third time):

- εf is a B-series $B(\beta(\theta), y)$:

$$\varepsilon \sum_{\mathbf{k}\in\mathbb{Z}^d} e^{i\mathbf{k}\cdot\theta} f_{\mathbf{k}}(y) = \sum_{u\in\mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \beta_u(\theta)\mathcal{F}_u(y),$$

with coefficients $\beta_u(\theta)$:

$$\beta_u(\theta) := \begin{cases} e^{i\mathbf{k}\cdot\theta} & \text{if } u = \textcircled{\mathbf{k}} \text{ for some } \mathbf{k}\in\mathbb{Z}^d, \\ 0 & \text{otherwise.} \end{cases}$$

- Hence (composition of maps becomes $*$) our initial value problem is:

$$\begin{aligned} \frac{d}{dt} B(\alpha(t), y_0) &= B(\alpha(t) * \beta(t\omega), y_0), \\ B(\alpha(0), y_0) &= B(\mathbb{1}, y_0). \end{aligned}$$

• In terms of the coefficients we have a universal formulation as an IVP in $\mathcal{G}$:

$$\frac{d}{dt}\alpha(t) = \alpha(t) * \beta(t\omega), \qquad \alpha(0) = \mathbb{1}.$$

• Use of the expression for $*$ leads to the recursive formulae for $\alpha_u(t)$ we found before.

(II.E) The coefficients γ

- Each $\alpha_u(t)$ in the expansion of $y(t)$ is of the form

$$\alpha_u(t) = \gamma_u(t, t\omega),$$

where γ is a polynomial in its fist argument and a trigonometric polynomial in the second.

Example. For $u = \binom{\mathbf{l}}{\mathbf{k}}$ with $\mathbf{k} = -\mathbf{l}$,

$$\alpha_u(t) = \frac{it}{\mathbf{l}\cdot\omega} + \frac{1 - e^{it\mathbf{l}\cdot\omega}}{(\mathbf{l}\cdot\omega)^2}$$

and therefore

$$\gamma_u(t,\theta) = \frac{it}{\mathbf{l}\cdot\omega} + \frac{1 - e^{i\mathbf{l}\cdot\theta}}{(\mathbf{l}\cdot\omega)^2}.$$

• $\gamma(t,\theta)$ satisfies $\gamma(0,\mathbf{0}) = \mathbb{1}$ and, since $(d/dt)\alpha(t) = (d/dt)\gamma(t,t\omega)$, the following transport equation:

$$\partial_t\gamma + \omega\cdot\nabla_\theta\gamma = \gamma * \beta(\theta).$$

This PDE replaces usual homological equation. Note is *autonomous* wrt t.

• From transport equation easy to obtain two important formulae:

$$\underline{\forall t, t' \in \mathbb{R}, \qquad \gamma(t',\mathbf{0}) * \gamma(t,\mathbf{0}) = \gamma(t'+t,\mathbf{0})}$$

(implied by trans. eq. being autonomous in t) and

$$\underline{\forall t \in \mathbb{R},\ \forall \theta \in \mathbb{T}^d, \qquad \gamma(t,\mathbf{0}) * \gamma(0,\theta) = \gamma(t,\theta).}$$

(II.F) Quasi-stroboscopic averaging

• From the first of these results, $\bar{\alpha}(t) := \gamma(t, \mathbf{0})$ form a one-parameter subgroup of $\mathcal{G}$ and hence satisfy an *autonomous* differential system:

$$\frac{d}{dt}\bar{\alpha}(t) = \bar{\alpha}(t) * \bar{\beta}, \quad \bar{\alpha}(0) = \mathbb{1}, \qquad \bar{\beta} := \frac{d}{dt}\bar{\alpha}(t)\Big|_{t=0}.$$

• From the second, if $\kappa(\theta) := \gamma(0, \theta)$, then $\alpha(t) = \bar{\alpha}(t) * \kappa(t\omega)$, so that κ provides a change of variables that maps the autonomous solution $\bar{\alpha}$ into the oscillatory solution α.

• Plug in the elementary differentials and obtain . . .

Theorem: The solution y may be written as $y(t) = U(Y(t), t\omega)$ where U is the change of variables (parameterized by $\theta \in \mathbb{T}^s$)

$$U(Y, \theta) := B(\kappa(\theta), Y) := Y + \sum_{u \in \mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \kappa_u(\theta)\, \mathcal{F}_u(Y)$$

and $Y(t)$ is the solution of the (averaged) autonomous initial value problem

$$\frac{d}{dt} Y = \varepsilon F(Y), \qquad Y(0) = y_0,$$

with

$$\varepsilon F(Y) := B(\bar{\beta}, Y) := \sum_{u \in \mathcal{T}} \frac{\varepsilon^{|u|}}{\sigma_u} \bar{\beta}_u\, \mathcal{F}_u(Y).$$

- Overview:

$$
\begin{array}{rcl}
\frac{d}{dt}y & = & \varepsilon f(y,t\omega) \\
f(y,\theta) & = & B(\beta(\theta),y) \\
 & \downarrow & \\
y(t) & = & B(\alpha(t),y_0) \\
\alpha(t) & = & \gamma(t,t\omega) \\
 & \downarrow & \\
y(t) & = & U(Y(t),t\omega) \\
U(Y,\theta) & = & B(\kappa(\theta),Y) \\
\kappa(\theta) & = & \gamma(0,\theta)
\end{array}
\qquad \rightarrow \qquad
\begin{array}{rcl}
\frac{d}{dt}Y & = & \varepsilon F(y) \\
F(Y) & = & B(\bar{\beta},Y) \\
 & \uparrow & \\
Y(t) & = & B(\bar{\alpha}(t),y_0) \\
\bar{\alpha}(t) & = & \gamma(t,\mathbf{0})
\end{array}
$$

• At time $t = 0$ the change of variables is the identity ($\kappa(\mathbf{0}) = \gamma(0, \mathbf{0}) = \mathbb{1}$).

• If $d = 1$, change of variables is also the identity at stroboscopic times $n(2\pi/\omega_1)$. Hence y and Y coincide at stroboscopic times. This is clearly not so if $d > 1$ and we refer to this technique as quasi-stroboscopic averaging.

• Other averaging strategies (e.g. zero-mean changes of variables) may be catered for. All possible averaged system result from q.s. system through a change of variables $Y = V(\widehat{Y})$.

III. REFINEMENTS

(III.A) Making the expansion more compact:

• By grouping elementary differentials together, the quasi-stroboscopically averaged system may be re-written in terms of iterated Lie-brackets:

$$\frac{d}{dt}Y = \sum_{r=1}^{\infty} \sum_{\mathbf{k}_1,\ldots,\mathbf{k}_r \in Z^d} \frac{\varepsilon^r}{r}\, \bar{\beta}_{\mathbf{k}_1\cdots\mathbf{k}_r}\, [[\cdots[[f_{\mathbf{k}_1}, f_{\mathbf{k}_2}], f_{\mathbf{k}_3}]\cdots], f_{\mathbf{k}_r}](Y).$$

• Here $\bar{\beta}_{\mathbf{k}_1\cdots\mathbf{k}_r}$ is the coefficient $\bar{\beta}_u$ corresponding to the un-ramified tree $u_{\mathbf{k}_1\cdots\mathbf{k}_r}$:

$$u_{\mathbf{k}} := \textcircled{\mathbf{k}}, \qquad u_{\mathbf{k}_1\cdots\mathbf{k}_r} := [u_{\mathbf{k}_1\cdots\mathbf{k}_{r-1}}]_{\mathbf{k}_r}.$$

- The $\bar{\beta}_{\mathbf{k}_1 \cdots \mathbf{k}_r}$ may be computed recursively (given ω):

$$
\begin{aligned}
\bar{\beta}_{\mathbf{k}} &= 0, \\
\bar{\beta}_{\mathbf{0}} &= 1, \\
\bar{\beta}_{\mathbf{0}^{r+1}} &= 0, \\
\bar{\beta}_{\mathbf{0}^r \mathbf{k}} &= \frac{i}{\mathbf{k} \cdot \omega} (\bar{\beta}_{\mathbf{0}^{r-1} \mathbf{k}} - \bar{\beta}_{\mathbf{0}^r}), \\
\bar{\beta}_{\mathbf{k} \mathbf{l}_1 \cdots \mathbf{l}_s} &= \frac{i}{\mathbf{k} \cdot \omega} (\bar{\beta}_{\mathbf{l}_1 \cdots \mathbf{l}_s} - \bar{\beta}_{(\mathbf{k}+\mathbf{l}_1) \mathbf{l}_2 \cdots \mathbf{l}_s}), \\
\bar{\beta}_{\mathbf{0}^r \mathbf{k} \mathbf{l}_1 \cdots \mathbf{l}_s} &= \frac{i}{\mathbf{k} \cdot \omega} (\bar{\beta}_{\mathbf{0}^{r-1} \mathbf{k} \mathbf{l}_1 \cdots \mathbf{l}_s} - \bar{\beta}_{\mathbf{0}^r (\mathbf{k}+\mathbf{l}_1) \mathbf{l}_2 \cdots \mathbf{l}_s}).
\end{aligned}
$$

- This recursion makes it possible to advantageously obtain *exponentially small error estimates*.

- Leading terms of the expansion: With an $\mathcal{O}(\varepsilon^4)$ remainder:

$$\frac{d}{dt}Y = \varepsilon F_1(Y) + \varepsilon^2 F_2(Y) + \varepsilon^3[F_3^+(Y) + F_3^-(Y)] + \mathcal{O}(\varepsilon^4),$$

where F_1, F_2, F_3^+ are, respectively, f_0

$$\sum_{\mathbf{k}>\mathbf{0}} \frac{i}{\mathbf{k}\cdot\omega}\left([f_{-\mathbf{k}}, f_{\mathbf{k}}] + [f_{\mathbf{k}} - f_{-\mathbf{k}}, f_{\mathbf{0}}]\right), \qquad \sum_{\mathbf{k}>\mathbf{0}} \frac{1}{(\mathbf{k}\cdot\omega)^2}$$

$$\left([f_{\mathbf{0}}, [f_{\mathbf{0}}, f_{\mathbf{k}}]] + [f_{\mathbf{k}}, [f_{\mathbf{k}}, f_{-\mathbf{k}}]] + \frac{1}{2}[f_{\mathbf{k}}, [f_{\mathbf{0}}, f_{\mathbf{k}}]] + 2[f_{-\mathbf{k}}, [f_{\mathbf{k}}, f_{\mathbf{0}}]]\right)$$

$$+ \sum_{\mathbf{k}>\mathbf{l}>\mathbf{0}} \frac{1}{(\mathbf{k}\cdot\omega)(\mathbf{l}\cdot\omega)}\left([f_{-\mathbf{l}}, [f_{\mathbf{k}}, f_{\mathbf{l}-\mathbf{k}}]] - [f_{\mathbf{l}}, [f_{\mathbf{k}}, f_{-\mathbf{k}-\mathbf{l}}]]\right),$$

and F_3^- is obtained from F_3^+ by replacing $(\mathbf{k}, \mathbf{l})$ by $(-\mathbf{k}, -\mathbf{l})$.

(III.B) Geometry:

• The possibility of re-writing in terms of Lie brackets ensures good geometric properties: for instance if the given system is canonical, with Hamiltonian function,

$$\varepsilon H(y,\theta) = \varepsilon \sum_{\mathbf{k}\in\mathbb{Z}^d} e^{i\mathbf{k}\cdot\theta} H_{\mathbf{k}}(y),$$

the q-s averaged system is canonical with Hamiltonian function:

$$\bar{H} = \sum_{r=1}^{\infty} \sum_{\mathbf{k}_1,\ldots,\mathbf{k}_r\in Z^d} \frac{\varepsilon^r}{r} \bar{\beta}_{\mathbf{k}_1\cdots\mathbf{k}_r} \{\{\cdots\{\{H_{\mathbf{k}_1}, H_{\mathbf{k}_2}\}, H_{\mathbf{k}_3}\}\cdots\}, H_{\mathbf{k}_r}\}.$$

The change of variables is also a canonical transformation.

(III.C) Near-integrable systems:

• As an application one may study Hamiltonians in $\mathbb{R}^D$ of the form

$$\mathcal{H}(x) := \sum_{j=1}^{d} \omega_j I_j(x) + \varepsilon K(x),$$

where the I_j are in involution, $d \leq D/2$ and generate 2π-periodic flows.

• It is possible to construct explicitly d functions $\widetilde{I}_j$ in involution, that are formal first integrals of both the given system and the averaged system.

• See paper `http://sanzserna.org/` for an application to find a new adiabatic invariant in a Fermi-Pasta-Ulam-type problem.