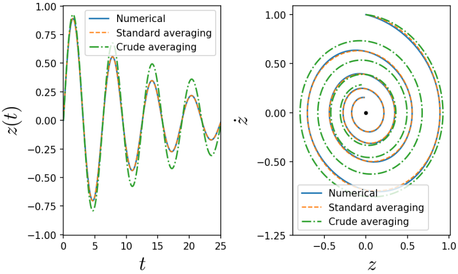
Averaging Methods

Iman Bayat

January 27, 2020



Definition 1

Consider the continuous vector field $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. We define the **local average** $\mathbf{f}_T$ of $\mathbf{f}$ by

$$\mathbf{f}_T(\boldsymbol{x}, t) = \frac{1}{T} \int_0^T \mathbf{f}(\boldsymbol{x}, t + s)\, \mathrm{d}s.$$

T is a parameter which can be chosen, and can be made ε-dependent if we wish. Whenever we estimate with respect to ε we shall require $\varepsilon T = o(1)$, where ε is a small parameter. So we may choose for instance $T = 1/\sqrt{\varepsilon}$ or $T = 1/|\varepsilon \log(\varepsilon)|$. Included is also $T = \mathcal{O}_\sharp(1)$ which we shall use for periodic $\mathbf{f}$. Note that the local average of a continuous vector field always exists.

Lemma 2

Consider the continuous vector field $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, *$T$-periodic in t. Then*

$$\mathbf{f}_T(\boldsymbol{x}, t) = \bar{\mathbf{f}}(\boldsymbol{x}) = \frac{1}{T}\int_0^T \mathbf{f}(\boldsymbol{x}, s)\,\mathrm{d}s.$$

We shall now introduce vector fields which can be averaged in a general sense. Since most applications are for differential equations we impose some additional regularity conditions on the vector field.

Definition 3

Consider the vector field $\mathbf{f}(\boldsymbol{x}, t)$ with $\mathbf{f} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, Lipschitz continuous in x on $D \subset \mathbb{R}^n$, $t \geq 0$; $\mathbf{f}$ continuous in t and $\boldsymbol{x}$ on $\mathbb{R}^+ \times D$. If the average

$$\bar{\mathbf{f}}(\boldsymbol{x}) = \lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbf{f}(\boldsymbol{x}, s)\,\mathrm{d}s.$$

exists and the limit is uniform in $\boldsymbol{x}$ on compact sets $K \subset D$, then $\mathbf{f}$ is called a **KBM-vector field** (KBM stands for Krylov, Bogoliubov and Mitropolsky).

Lemma 4

Consider the Lipschitz continuous map $\boldsymbol{x} : \mathbb{R} \to \mathbb{R}^n$ *with Lipschitz constant* $\lambda_{\boldsymbol{x}}$*, then*

$$\|\mathbf{x}(t) - \mathbf{x}_T(t)\| \leq \frac{1}{2}\lambda_{\mathbf{x}} T.$$

Proof.

One has

$$\|\mathbf{x}(t) - \mathbf{x}_T(t)\| = \left\| \frac{1}{T} \int_0^T (\mathbf{x}(t) - \mathbf{x}(t+s))\, \mathrm{d}s \right\| \leq \frac{1}{T} \int_0^T \lambda_{\mathbf{x}} s \, \mathrm{d}s = \frac{1}{2}\lambda_{\mathbf{x}} T,$$

and this gives the desired estimate. □

Corollary 5

Let $\mathbf{x}(t)$ *be a solution of the equation*

$$\dot{\boldsymbol{x}} = \varepsilon \mathbf{f}^1(\boldsymbol{x}, t), \quad t \geq 0, \quad \boldsymbol{x} \in D \subset \mathbb{R}^n.$$

Let

$$M = \sup_{\boldsymbol{x} \in D} \sup_{0 \leq \varepsilon t \leq L} \|\mathbf{f}^1(\boldsymbol{x}, t)\| < \infty.$$

Then $\|\mathbf{x}(t) - \mathbf{x}_T(t)\| \leq \frac{1}{2}\varepsilon M T$. *(since* $\lambda_{\mathbf{x}} = \varepsilon M$*)*

The preceding lemmas enable us to compare solutions of two differential equations:

Lemma 6

Consider the initial value problem

$$\dot{\boldsymbol{x}} = \varepsilon \mathbf{f}^1(\boldsymbol{x}, t), \quad \mathbf{x}(0) = \boldsymbol{a},$$

with $\mathbf{f}^1 : \mathbb{R}^n \times \mathbb{R}$ *Lipschitz continuous in* $\boldsymbol{x}$ *on* $D \subset \mathbb{R}^n$, t *on the time scale* $1/\varepsilon$; $\mathbf{f}^1$ *continuous in* t *and* $\boldsymbol{x}$. *If* $\boldsymbol{y}$ *is the solution of*

$$\dot{\boldsymbol{y}} = \varepsilon \mathbf{f}^1_T(\boldsymbol{y}, t), \quad \mathbf{y}(0) = \boldsymbol{a},$$

then $\mathbf{x}(t) = \mathbf{y}(t) + \mathcal{O}(\varepsilon T)$ *on the time scale* $1/\varepsilon$.

Proof.

Writing the differential equation as an integral equation, we see that

$$\mathbf{x}(t) = \boldsymbol{a} + \varepsilon \int_0^t \mathbf{f}^1(\mathbf{x}(s), s)\,\mathrm{d}s.$$

With Corollary 5 and Lemma **??** we obtain

$$\begin{aligned}
&\left\|\mathbf{x}(t) - \boldsymbol{a} - \varepsilon \int_0^t \mathbf{f}_T^1(\mathbf{x}(s), s)\,\mathrm{d}s\right\| \\
&\leq \|\mathbf{x}(t) - \mathbf{x}_T(t)\| + \left\|\mathbf{x}_T(t) - \boldsymbol{a} - \varepsilon \int_0^t \mathbf{f}_T^1(\mathbf{x}(s), s)\,\mathrm{d}s\right\| \\
&\leq \varepsilon M T (1 + \frac{1}{2}\lambda_{\mathbf{f}^1} L).
\end{aligned}$$

Proof.

It follows that $\mathbf{x}(t) = \boldsymbol{a} + \varepsilon \int_0^t \mathbf{f}_T(\mathbf{x}(s), s)\,\mathrm{d}s + \mathcal{O}(\varepsilon T)$. Since

$$\mathbf{y}(t) = \boldsymbol{a} + \varepsilon \int_0^t \mathbf{f}_T^1(\mathbf{y}(s), s)\,\mathrm{d}s,$$

we have

$$\mathbf{x}(t) - \mathbf{y}(t) = \varepsilon \int_0^t [\mathbf{f}_T^1(\mathbf{x}(s), s) - \mathbf{f}_T^1(\mathbf{y}(s), s)]\,\mathrm{d}s + \mathcal{O}(\varepsilon T),$$

and because of the Lipschitz continuity of $\mathbf{f}_T^1$ (inherited from $\mathbf{f}^1$)

$$\|\mathbf{x}(t) - \mathbf{y}(t)\| \leq \varepsilon \int_0^t \lambda_{\mathbf{f}^1} \|\mathbf{x}(s) - \mathbf{y}(t)\|\,\mathrm{d}s + \mathcal{O}(\varepsilon T).$$

The Gronwall Lemma yields

$$\|\mathbf{x}(t) - \mathbf{y}(t)\| = \mathcal{O}(\varepsilon T e^{\varepsilon \lambda_{\mathbf{f}^1} t}),$$

from which the lemma follows. □

Theorem 7 (Periodic averaging)

Consider the initial value problems

$$\dot{\boldsymbol{x}} = \varepsilon \mathbf{f}^1(\boldsymbol{x}, t), \quad \mathbf{x}(0) = \boldsymbol{a},$$

with $\mathbf{f}^1 : \mathbb{R}^{n+1} \to \mathbb{R}^n$ *and*

$$\dot{\boldsymbol{z}} = \varepsilon \overline{\mathbf{f}}^1(\boldsymbol{z}), \quad \mathbf{z}(0) = \boldsymbol{a},$$

where $\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{a} \in D \subset \mathbb{R}^n$, $t \in [0, \infty)$, $\varepsilon \in (0, \varepsilon_0]$. *Suppose*

1. $\mathbf{f}^1$ *has period* T*;*
2. $\mathbf{f}^1$ *is Lipschitz continuous in* $\boldsymbol{x}$ *on* $D \subset \mathbb{R}^n$, $t \geq 0$, *continuous in* t *and* $\boldsymbol{x}$ *on* $\mathbb{R}^+ \times D$ *and with average* $\overline{\mathbf{f}}^1$*;*
3. $\mathbf{z}(t)$ *belongs to an interior subset of* D *on the time scale* $1/\varepsilon$*;*

then

$$\mathbf{x}(t) - \mathbf{z}(t) = \mathcal{O}(\varepsilon)$$

as $\varepsilon \downarrow 0$ *on the time scale* $1/\varepsilon$.

To prove the fundamental theorem of general averaging we need a few more results.

Lemma 8

If $\mathbf{f}^1$ *is a KBM-vector field and assuming* $\varepsilon T = o(1)$ *as* $\varepsilon \downarrow 0$*, then on the time scale* $1/\varepsilon$ *one has*

$$\mathbf{f}^1_T(\boldsymbol{x}, t) = \overline{\mathbf{f}}^1(\boldsymbol{x}) + \mathcal{O}(\delta_1(\varepsilon)/(\varepsilon T)),$$

where

$$\delta_1(\varepsilon) = \sup_{\boldsymbol{x}\in D} \sup_{t\in[0,L/\varepsilon)} \varepsilon \left\| \int_0^t [\mathbf{f}^1(\boldsymbol{x}, s) - \overline{\mathbf{f}}^1(\boldsymbol{x})] \,\mathrm{d}s \right\|.$$

Remark

We call $\delta_1(\varepsilon)$ the **order function of** $\mathbf{f}^1$. In the periodic case $\delta_1(\varepsilon) = \varepsilon$.

Lemma 9

Let $\mathbf{y}$ *be the solution of the initial value problem*

$$\dot{\boldsymbol{y}} = \varepsilon \mathbf{f}^1_T(\boldsymbol{y}, t), \quad \mathbf{y}(0) = \boldsymbol{a}.$$

We suppose $\mathbf{f}^1$ *is a KBM-vector field with order function* $\delta_1(\varepsilon)$ *; let* $\mathbf{z}$ *be the solution of the initial value problem*

$$\dot{\boldsymbol{z}} = \varepsilon \overline{\mathbf{f}}^1(\boldsymbol{z}), \quad \mathbf{z}(0) = \boldsymbol{a}.$$

Then

$$\mathbf{y}(t) = \mathbf{z}(t) + \mathcal{O}(\delta_1(\varepsilon)/(\varepsilon T)),$$

with t *on the time scale* $1/\varepsilon$.

We are now able to prove the general averaging theorem:

Theorem 10 (general averaging)

Consider the initial value problems

$$\dot{\boldsymbol{x}} = \varepsilon \mathbf{f}^1(\boldsymbol{x}, t), \quad \mathbf{x}(0) = \boldsymbol{a},$$

with $\mathbf{f}^1 : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ *and*

$$\dot{\boldsymbol{z}} = \varepsilon \overline{\mathbf{f}}^1(\boldsymbol{z}), \quad \mathbf{z}(0) = \boldsymbol{a},$$

where

$$\overline{\mathbf{f}}^1(\boldsymbol{x}) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \mathbf{f}^1(\boldsymbol{x}, t) \, \mathrm{d}t,$$

and $\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{a} \in D \subset \mathbb{R}^n$, $t \in [0, \infty)$, $\varepsilon \in (0, \varepsilon_0]$.

Theorem 10 (general averaging)

Suppose

1. $\mathbf{f}^1$ *is a KBM-vector field with average* $\overline{\mathbf{f}}^1$ *and order function* $\delta_1(\varepsilon)$;
2. $\mathbf{z}(t)$ *belongs to an interior subset of* D *on the time scale* $1/\varepsilon$;

then

$$\mathbf{x}(t) - \mathbf{z}(t) = \mathcal{O}(\sqrt{\delta_1(\varepsilon)})$$

as $\varepsilon \downarrow 0$ *on the time scale* $1/\varepsilon$.

Proof.

Applying Lemmas 6 and 9, using the triangle inequality, we have on the time scale $1/\varepsilon$:

$$\mathbf{x}(t) = \mathbf{z}(t) + \mathcal{O}(\varepsilon T) + \mathcal{O}(\delta_1(\varepsilon)/(\varepsilon T)).$$

The errors are of the same order of magnitude if

$$\varepsilon^2 T^2 = \delta_1(\varepsilon),$$

so that

$$\mathbf{x}(t) = \mathbf{z}(t) + \mathcal{O}(\sqrt{\delta_1(\varepsilon)}),$$

if we let $T = \sqrt{\delta_1(\varepsilon)}/\varepsilon$. □

Consider the equation

$$\ddot{x} + \varepsilon(2 - F(t))\dot{x} + x = 0,$$

with initial values given at $t = 0 : x(0) = r_0$, $\dot{x}(0) = 0$. $F(t)$ is a continuous function, monotonically decreasing towards zero for $t \to \infty$ with $F(0) = 1$. So the problem is simple: we start with an oscillator with damping coefficient ε, we end up (in the limit for $t \to \infty$) with an oscillator with damping coefficient 2ε. We shall show that on the time scale $1/\varepsilon$ the system behaves approximately as if it has the limiting damping coefficient 2ε, which seems an interesting result. To obtain the standard form, transform $(x, \dot{x}) \mapsto (r, \phi)$ by

$$x = r\cos(t + \phi), \quad \dot{x} = -r\sin(t + \phi).$$

We obtain

$$\dot{r} = \varepsilon r \sin^2(t + \phi)(-2 + F(t)), \quad r(0) = r_0,$$
$$\dot{\phi} = \varepsilon \sin(t + \phi)\cos(t + \phi)(-2 + F(t)), \quad \phi(0) = 0.$$

Averaging produces

$$\dot{\tilde{r}} = -\varepsilon\tilde{r}, \quad \dot{\tilde{\phi}} = 0,$$

so that

$$x(t) = r_0 e^{-\varepsilon t}\cos(t) + \mathcal{O}(\sqrt{\delta_1(\varepsilon)}),$$
$$\dot{x}(t) = -r_0 e^{-\varepsilon t}\sin(t) + \mathcal{O}(\sqrt{\delta_1(\varepsilon)}).$$

To estimate δ_1 we note that $x = \dot{x} = 0$ is a globally stable attractor (one can use the Lyapunov function $\frac{1}{2}(x^2 + \dot{x}^2)$ to show this if the mechanics of the problem is not already convincing enough). So the order of magnitude of δ_1 is determined by

$$\sup_{D} \sup_{t\in[0,\frac{L}{\varepsilon})} \varepsilon \left| \int_0^t [\sin^2(s+\phi)(-2+F(s))+1]\,\mathrm{d}s \right|$$

and

$$\sup_{D} \sup_{t\in[0,\frac{L}{\varepsilon})} \varepsilon \left| \int_0^t [\sin(s+\phi)\cos(s+\phi)(-2+F(s))]\,\mathrm{d}s \right|.$$

The second integral is bounded for all t so this contributes $\mathcal{O}(\varepsilon)$. The same holds for the part

$$\int_0^t (-2\sin^2(s+\phi)+1)\,\mathrm{d}s.$$

To estimate

$$\int_0^t F(s)\sin^2(s+\phi)\,\mathrm{d}s$$

we have to make an assumption about F. For instance if F decreases exponentially with time we have $\delta_1(\varepsilon) = \mathcal{O}(\varepsilon)$ and an approximation with error $\mathcal{O}(\sqrt{\varepsilon})$. If $F \approx t^{-s}(0 < s < 1)$ we have $\delta_1(\varepsilon) = \mathcal{O}(\varepsilon^s)$ and an approximation with error $\mathcal{O}(\varepsilon^{\frac{s}{2}})$. If $F(t) = (1+t)^{-1}$ we have $\delta_1(\varepsilon) = \mathcal{O}(\varepsilon|\log(\varepsilon)|)$. If $\delta_1(\varepsilon)$ is not $o(1)$, then we need to adapt the average in order to apply the theory. We remark finally that to describe the dependence of the oscillator on the initial damping we clearly need a different order of approximation.