Applied Mathematics Letters 178 (2026) 109910

Contents lists available at ScienceDirect

Applied Mathematics Letters

journal homepage: www.elsevier.com/locate/aml

Regular article

A note on the averaging principle for ordinary differential equations depending on the slow time☆

Carmen Núñez[1], Rafael Obaya[1], Jorge Rodríguez*,[1]

Departamento de Matemática Aplicada, Universidad de Valladolid, Paseo Prado de la Magdalena 3-5, 47011 Valladolid, Spain

ARTICLE INFO

MSC:
34C29
34E13
37B55

Keywords:
Averaging theory
Nonautonomous dynamical systems
Multiscale ordinary differential equations

ABSTRACT

The work presents a "doubly" nonautonomous version of the averaging principle, applicable to equations that depend on a small parameter ε and on (fast) time τ, but also on slow time $t = \varepsilon\tau$. The objectives are to establish optimal conditions on the dependence of the coefficients of the equations on t under which the averaging principle can be extended and to provide good estimates of the distance between the solutions of the initial equation and those of the averaged equation, always with τ varying in intervals of length proportional to $1/\varepsilon$. The applicability of these results is based on the fact that the estimates obtained are uniform with respect to the initial time at which the solutions of both equations coincide.

1. Introduction

During the last decades, the theory of averaging has demonstrated a high degree of applicability in the study of the asymptotic behavior of the orbits of dynamic systems. The idea is to measure the distance between the solutions of a time-dependent n-dimensional ordinary differential equation written in standard form, $dx/d\tau = \varepsilon f(\tau, x)$, and those of the averaged equation, $dz/d\tau = \varepsilon \hat{f}(z)$, where $\hat{f}$ is constructed as the mean in time of f. Here, ε is a small parameter, which causes the independent variable τ to be referred to as fast time. The averaging principle establishes estimates for the distance between the solutions of both equations with the same initial data at intervals of length proportional to $1/\varepsilon$. The works of Bogoliubov and Mitropolsky [1], Hale [2], Arnold [3], and Mitropolsky [4] contain pioneering results of the averaging theory with notable examples and applications; the books by Sanders and Verhulst [5] and Sanders et al. [6] provide a comprehensive analysis of this theory in the context of periodic and quasi-periodic differential equations; and the more recent article Artstein [7] introduces the notion of rate of averaging to deduce accurate estimates for the distance between solutions, reaching conclusions which have been extended and applied in Bright [8] and Artstein [9].

Some notable applications in oscillation and climatology theories, among others, justify the interest in including a slow time in the above equations. Therefore, in this note we consider equations that also depend explicitly on the slow time $t = \varepsilon\tau$. That is, they take the form $dx/d\tau = \varepsilon f(\tau, \varepsilon\tau, x) = \varepsilon f(\tau, t, x)$ and $dz/d\tau = \varepsilon \hat{f}(\varepsilon\tau, z) = \varepsilon \hat{f}(t, x)$. This case has been analyzed in [4–6,9] considering t as a new dependent variable (adding the equation $dt/d\tau = \varepsilon$), and adding to the classical hypotheses the Lipschitz variation of f with respect to t: this allows the theory previously developed to be applied. However, a thorough analysis of the averaged equation should consider it as a nonautonomous equation, with t as an independent variable: such an approach would

☆ All the authors were supported by Ministerio de Ciencia, Innovación y Universidades (Spain) under project PID2024-156691NB-I00. J. Rodríguez was also supported by Ministerio de Ciencia, Innovación y Universidades (Spain) under programme FPU24/00913.

* Corresponding author.
E-mail addresses: carmen.nunez@uva.es (C. Núñez), rafael.obaya@uva.es (R. Obaya), jorge.rodriguez@uva.es (J. Rodríguez).
[1] The three authors are members of IMUVA, Universidad de Valladolid.

https://doi.org/10.1016/j.aml.2026.109910
Received 20 January 2026; Received in revised form 2 March 2026; Accepted 2 March 2026
Available online 4 March 2026

allow the use of dynamic methods from nonautonomous theory to deduce relevant properties of the averaged equation, and hence of the initial one. The nonautonomous framework should avoid the condition of Lipschitz variation of f in t, which is generally an unnecessarily restrictive assumption.

The objective of the current work is to provide appropriate hypotheses that allow for the best application of standard methods in the theories of averaging and of nonautonomous dynamical systems [10]. We prove the averaging principle when the families of maps $\{\mathbb{R}_+\times\mathcal{K}\to\mathbb{R}^n\,,\ (t,x)\mapsto f(\tau,t,x)\mid \tau\in\mathbb{R}_+\}$ are equicontinuous for each compact subset $\mathcal{K}\subset\mathbb{R}^n$, and provide quantitative estimates for the difference between the solutions of the original and the averaged equations on intervals $[\tau^*,\ \tau^*+c/\varepsilon]$ for each fixed $c>0$. In addition, these estimates are uniform on $\tau^*\ge 0$, which is a matter of high interest in applied problems. We also show that the weaker hypothesis of equicontinuity of the families $\{[0,c]\times\mathcal{K}\to\mathbb{R}^n\,,\ (t,x)\mapsto f(\tau,t,x)\mid\tau\in\mathbb{R}_+\}$ for each $c>0$ and each compact subset $\mathcal{K}\subset\mathbb{R}^n$ imply the validity of the above estimates on intervals $[0,c/\varepsilon]$: this weaker result also provides an extension on the well-known autonomous averaging results to the nonautonomous setting. We complete this note with a simple example showing that, when the hypotheses of equicontinuity of f fails, also the averaging principle fails: the maximum distance between the values of the solutions of the initial and the averaged equation in intervals of length proportional to $1/\varepsilon$ does not converge to 0 as ε decreases to 0.

2. Averaging results

Let $x_\varepsilon(\tau;\tau^*,x^*)$ and $z_\varepsilon(\tau;\tau^*,x^*)$ be the respective solutions with value x^* at time τ^* of equations

$$\frac{dx}{d\tau}=\varepsilon f(\tau,\varepsilon\tau,x)\qquad\text{and}\qquad\frac{dz}{d\tau}=\varepsilon\hat f(\varepsilon\tau,z)\,,\tag{2.1}$$

where $\hat f$ is the τ-mean function of f (below defined). Our purpose is to bound the difference $|x_\varepsilon(\tau;\tau^*,x^*)-z_\varepsilon(\tau;\tau^*,x^*)|$ for τ varying in time intervals of length proportional to ε^{-1}. We will refer to τ as the *fast* time in contrast to the *slow* time variable $t:=\varepsilon\tau$. This change of variables provides the equations

$$\frac{dx}{dt}=f(t/\varepsilon,t,x)\qquad\text{and}\qquad\frac{dz}{dt}=\hat f(t,z)\,,\tag{2.2}$$

and, if $\tilde x_\varepsilon(t;t^*,x^*)$ and $\tilde z(t;t^*,x^*)$ are the respective solutions with value x^* at time t^*, then

$$x_\varepsilon(\tau;\tau^*,x^*)=\tilde x_\varepsilon(\varepsilon\tau;\varepsilon\tau^*,x^*)\quad\text{and}\quad z_\varepsilon(\tau;\tau^*,x^*)=\tilde z(\varepsilon\tau;\varepsilon\tau^*,x^*).\tag{2.3}$$

We denote $\mathbb{R}_+:=[0,\infty)$. The distance in $\mathbb{R}_+$ and $\mathbb{R}^n$ is that of the Euclidean norm $|\cdot|$, $\mathrm{d}((t_1,x_1),(t_2,x_2))=\max\{|t_1-t_2|,|x_1-x_2|\}$ in $\mathbb{R}_+\times\mathbb{R}^n$, and $\mathrm{d}((\tau_1,t_1,x_1),(\tau_2,t_2,x_2))=\max\{|\tau_1-\tau_2|,\ |t_1-t_2|,|x_1-x_2|\}$ in $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^n$.

2.1. Uniform averaging results

Recall that a *modulus of continuity* is a nondecreasing map $\theta:\mathbb{R}_+\to\mathbb{R}_+\cup\{\infty\}$ with $\lim_{\delta\to0^+}\theta(\delta)=\theta(0)=0$.

Hypotheses 2.1. The map $f:\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}^n$ is continuous and, for every $\mathcal{K}\subset\mathbb{R}^n$ compact: there exists $M_\mathcal{K}\ge0$ such that $|f|\le M_\mathcal{K}$ on $\mathbb{R}_+\times\mathbb{R}_+\times\mathcal{K}$; there exist a modulus of continuity $\theta_\mathcal{K}$ and $L_\mathcal{K}>0$ such that $|f(\tau,t_1,x_1)-f(\tau,t_2,x_2)|\le\theta_\mathcal{K}(|t_1-t_2|)+L_\mathcal{K}|x_1-x_2|$ for all $\tau\in\mathbb{R}_+$, $t_1,t_2\in\mathbb{R}_+$ and $x_1,x_2\in\mathcal{K}$; there exists $\hat f(t,x):=\lim_{T\to\infty}(1/T)\int_0^T f(\sigma,t,x)\,d\sigma$ for all $(t,x)\in\mathbb{R}\times\mathbb{R}^n$; and $\hat f(t,x)=\lim_{T\to\infty}(1/T)\int_\tau^{T+\tau}f(\sigma,t,x)\,d\sigma$ uniformly on $(\tau,t,x)\in\mathbb{R}_+\times\mathbb{R}_+\times\mathcal{K}$.

Remark 2.2. It is easy to deduce that $|\hat f(t_1,x_1)|\le M_\mathcal{K}$ and $|\hat f(t_1,x_1)-\hat f(t_2,x_2)|\le\theta_\mathcal{K}(|t_1-t_2|)+L_\mathcal{K}|x_1-x_2|$ for all $t_1,t_2\in\mathbb{R}_+$ and $x_1,x_2\in\mathcal{K}$. In particular, $\hat f$ is bounded and uniformly continuous on $\mathbb{R}^+\times\mathcal{K}$.

Definition 2.3. Assume Hypotheses 2.1, and fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact. The *uniform static rate of convergence of f on* $[0,c]\times\mathcal{K}$ is

$$\hat\delta^S_{c,\mathcal{K}}(\varepsilon):=\sup_{(\tau^*,t,x)\in\mathbb{R}_+\times\mathbb{R}_+\times\mathcal{K},\ T\in[0,c/\varepsilon]}\varepsilon\left|\int_{\tau^*}^{\tau^*+T}(f(\sigma,t,x)-\hat f(t,x))\,d\sigma\right|,$$

and the *uniform dynamic rate of convergence of f on* $[0,c]\times\mathcal{K}$ is

$$\hat\delta^D_{c,\mathcal{K}}(\varepsilon):=\sup_{(\tau^*,x)\in\mathbb{R}_+\times\mathcal{K},\ T\in[0,c/\varepsilon]}\varepsilon\left|\int_{\tau^*}^{\tau^*+T}(f(\sigma,\varepsilon\sigma,x)-\hat f(\varepsilon\sigma,x))\,d\sigma\right|.$$

Proposition 2.4. *Assume Hypotheses 2.1, and fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact. Then,*

(i) $\lim_{\varepsilon\to0^+}\hat\delta^S_{c,\mathcal{K}}(\varepsilon)=0$, (ii) $\lim_{\varepsilon\to0^+}\hat\delta^D_{c,\mathcal{K}}(\varepsilon)=0$.

Proof. (i) We take $\rho>0$. The uniform convergence of Hypotheses 2.1 ensures the existence of $T_0=T_0(\mathcal{K},\rho)>0$ such that $\sup_{(\tau^*,t,x)\in\mathbb{R}_+\times\mathbb{R}_+\times\mathcal{K}}\varepsilon T\left|(1/T)\int_{\tau^*}^{\tau^*+T}f(\sigma,t,x)\,d\sigma-\hat f(t,x)\right|<\rho$ for every $\varepsilon>0$ and $T\in[T_0,c/\varepsilon]$. On the other hand, $\sup_{(\tau^*,t,x)\in\mathbb{R}_+\times\mathbb{R}_+\times\mathcal{K}}\varepsilon T$

$\left|(1/T)\int_{\tau^*}^{\tau^*+T} f(\sigma,t,x)\,d\sigma - \hat f(t,x)\right| < \rho$ for every $T\in[0,T_0]$ if $0<\varepsilon<\rho/(2T_0M_{\mathcal K})$, where $M_{\mathcal K}\geq |f|$ on $\mathbb R_+\times\mathbb R_+\times\mathcal K$. So, $\hat\delta^S_{c,\mathcal K}(\varepsilon)<\rho$ if $0<\varepsilon<\rho/(2T_0M_{\mathcal K})$.

(ii) For simplicity, we call $\delta_\varepsilon := \hat\delta^S_{c,\mathcal K}(\varepsilon)$. We take $\varepsilon>0$ small enough to get $\sqrt{\delta_\varepsilon}\in[0,c]$ and note that

$$\delta_\varepsilon = \sup_{(t^*,t,x)\in\mathbb R_+\times\mathbb R_+\times\mathcal K,\,T\in[0,c]} \left|\int_{t^*}^{t^*+T}(f(s/\varepsilon,t,x)-\hat f(t,x))\,ds\right| \geq \left|\int_{t^*}^{t^*+\sqrt{\delta_\varepsilon}}(f(s/\varepsilon,t,x)-\hat f(t,x))\,ds\right| \tag{2.4}$$

for all $(t^*,t,x)\in\mathbb R_+\times\mathbb R_+\times\mathcal K$, and that $\hat\delta^D_{c,\mathcal K}(\varepsilon)=\sup_{(t^*,x)\in\mathbb R_+\times\mathcal K,\,T\in[0,c]}\left|\int_{t^*}^{t^*+T}(f(s/\varepsilon,s,x)-\hat f(s,x))\,ds\right|$. We fix $t_0\in\mathbb R_+$, $x\in\mathcal K$ and $T\in[0,c]$, and divide the interval $[t_0,t_0+T]$ in the subintervals $\mathcal I_0 := [t_0,\,t_0+\sqrt{\delta_\varepsilon})$, $\mathcal I_1 := [t_0+\sqrt{\delta_\varepsilon},\,t_0+2\sqrt{\delta_\varepsilon}),\ldots,\mathcal I_{m_\varepsilon} := [t_0+m_\varepsilon\sqrt{\delta_\varepsilon},\,t_0+T]$, where the length of $\mathcal I_{m_\varepsilon}$ is at most $\sqrt{\delta_\varepsilon}$. Let $t_j := \inf\mathcal I_j$. Let $M_{\mathcal K}$ be a bound for $|f|$ on $\mathbb R_+\times\mathbb R_+\times\mathcal K$. Then, by Hypotheses 2.1, Remark 2.2 and (2.4),

$$\begin{aligned}\left|\int_{t_0}^{t_0+T}(f(s/\varepsilon,s,x)-\hat f(s,x))\,ds\right| &\leq \sum_{j=0}^{m_\varepsilon}\int_{\mathcal I_j}|f(s/\varepsilon,s,x)-f(s/\varepsilon,t_j,x)|\,ds\\ &\quad+\sum_{j=0}^{m_\varepsilon}\left|\int_{\mathcal I_j}(f(s/\varepsilon,t_j,x)-\hat f(t_j,x))\,ds\right|+\sum_{j=0}^{m_\varepsilon}\int_{\mathcal I_j}|\hat f(t_j,x)-\hat f(s,x)|\,ds\\ &\leq \sum_{j=0}^{m_\varepsilon}\int_{\mathcal I_j}\theta_{\mathcal K}(|s-t_j|)\,ds+\sum_{j=0}^{m_\varepsilon-1}\delta_\varepsilon+2M_{\mathcal K}\sqrt{\delta_\varepsilon}+\sum_{j=0}^{m_\varepsilon}\int_{\mathcal I_j}\theta_{\mathcal K}(|s-t_j|)\,ds\,.\end{aligned}$$

As $\theta_{\mathcal K}$ is nondecreasing, the first and last terms are bounded by $\theta_{\mathcal K}(\sqrt{\delta_\varepsilon})\sum_{j=0}^{m_\varepsilon}\int_{\mathcal I_j}d\sigma = T\,\theta_{\mathcal K}(\sqrt{\delta_\varepsilon})\leq c\,\theta_{\mathcal K}(\sqrt{\delta_\varepsilon})$, while the second term can be bounded as $\sum_{j=0}^{m_\varepsilon-1}\delta_\varepsilon \leq \sqrt{\delta_\varepsilon}\sum_{j=0}^{m_\varepsilon-1}\sqrt{\delta_\varepsilon}\leq T\sqrt{\delta_\varepsilon}\leq c\sqrt{\delta_\varepsilon}$. Altogether, we get $\left|\int_{t_0}^{t_0+T}(f(s/\varepsilon,s,x)-\hat f(s,x))\,ds\right| \leq 2c\,\theta_{\mathcal K}(\sqrt{\delta_\varepsilon})+(c+2M_{\mathcal K})\sqrt{\delta_\varepsilon}$ for all $(t_0,x)\in\mathbb R_+\times\mathcal K$ and $T\in[0,c]$, which combined with (i) and with $\lim_{\varepsilon\to0^+}\theta_{\mathcal K}(\sqrt{\delta_\varepsilon})=0$ ensures (ii). □

Definition 2.5. Assume Hypotheses 2.1, and fix $c>0$ and $\mathcal X\subseteq\mathbb R^n$. A pair of nonnegative functions $(\hat\Delta(\varepsilon),\hat\eta(\varepsilon))$ is a *uniform rate of averaging of f on* $[0,c]\times\mathcal X$ if $\lim_{\varepsilon\to0^+}\hat\Delta(\varepsilon)=\lim_{\varepsilon\to0^+}\hat\eta(\varepsilon)=0$ and

$$\frac{\varepsilon}{\hat\Delta(\varepsilon)}\left|\int_\tau^{\tau+\hat\Delta(\varepsilon)/\varepsilon}(f(\sigma,\varepsilon\sigma,x)-\hat f(\varepsilon\sigma,x))\,d\sigma\right|\leq\hat\eta(\varepsilon) \tag{2.5}$$

for every $(\tau,x)\in\mathbb R_+\times\mathcal X$ and every $\varepsilon\in(0,\infty)$ with $\hat\Delta(\varepsilon)\in[0,c]$.

Proposition 2.6. *Assume Hypotheses 2.1, and fix $c>0$ and $\mathcal K\subset\mathbb R^n$ compact. For any $\alpha\in(0,1)$, the pair $\big((\hat\delta^D_{c,\mathcal K}(\varepsilon))^\alpha,\,(\hat\delta^D_{c,\mathcal K}(\varepsilon))^{1-\alpha}\big)$ is a uniform rate of averaging of f on $[0,c]\times\mathcal K$.*

Proof. Proposition 2.4(ii) ensures that both maps tend to 0 as $\varepsilon\to0^+$. The definition of $\hat\delta^D_{c,\mathcal K}(\varepsilon)$ yields

$$\varepsilon\left|\int_\tau^{\tau+(\hat\delta^D_{c,\mathcal K}(\varepsilon))^\alpha/\varepsilon}(f(\sigma,\varepsilon\sigma,x)-\hat f(\varepsilon\sigma,x))\,d\sigma\right|\leq\hat\delta^D_{c,\mathcal K}(\varepsilon)$$

for every $(\tau,x)\in\mathbb R_+\times\mathcal K$ if $(\hat\delta^D_{c,\mathcal K}(\varepsilon))^\alpha\leq c$. The result follows by dividing by $(\hat\delta^D_{c,\mathcal K}(\varepsilon))^\alpha$ at both sides. □

Let us fix $c>0$, and take $(t^*,x^*)\in\mathbb R_+\times\mathbb R^n$. The proofs of Theorems 2.8 and 2.16 require to consider the maps $[t^*,t^*+c]\to\mathbb R^n$, $t\mapsto\tilde x_\varepsilon(t;t^*,x^*)$, $\tilde z(t;t^*,x^*)$, which solve Eqs. (2.2) on $[t^*,t^*+c]$ with value x^* at t^*, as fixed points of the maps $S,\hat S: C([t^*,t^*+c],\mathbb R^n)\to C([t^*,t^*+c],\mathbb R^n)$ given by

$$S(\tilde x)(t) := x^*+\int_{t^*}^t f(s/\varepsilon,s,\tilde x(s))\,ds \qquad\text{and}\qquad \hat S(\tilde x)(t) := x^*+\int_{t^*}^t\hat f(s,\tilde x(s))\,ds\,. \tag{2.6}$$

The next technical lemma will also be used. It reproduces [7, Lemma 2.2]. A short and complete proof is given for the reader's convenience.

Lemma 2.7. *Let us fix $c>0$, take $t^*\geq0$ and $x^*\in\mathbb R^n$, and define $\|\tilde x\|_\infty := \max_{t\in[t^*,t^*+c]}|\tilde x(t)|$ and $\|\tilde x\|_{Li} := \max_{t\in[t^*,t^*+c]}|\tilde x(t)|\,e^{-(t-t^*)L_{\mathcal K}}$ for $\tilde x\in C([t^*,t^*+c],\mathbb R^n)$. Then, $\|\tilde x\|_{Li}\leq\|\tilde x\|_\infty\leq e^{cL_{\mathcal K}}\|\tilde x\|_{Li}$. In addition, if $\mathcal K\subset\mathbb R^n$ is compact and if $f:\mathbb R_+\times\mathbb R_+\times\mathcal K\to\mathbb R^n$ is Lipschitz in $x\in\mathcal K$ with constant $L_{\mathcal K}>0$, then the map S of (2.6) satisfies $\|S(\tilde x_1)-S(\tilde x_2)\|_{Li}\leq(1-e^{-cL_{\mathcal K}})\|\tilde x_1-\tilde x_2\|_{Li}$ for all $\tilde x_1,\,\tilde x_2\in C([t^*,t^*+c],\mathcal K)$.*

Proof. Since $|\tilde x(t)|\,e^{-(t-t^*)L_{\mathcal K}}\leq|\tilde x(t)|\leq|\tilde x(t)|\,e^{(c-(t-t^*))L_{\mathcal K}}$ for every $t\in[t^*,t^*+c]$, $\|\tilde x\|_{\rm Li}\leq\|\tilde x\|_\infty\leq e^{cL_{\mathcal K}}\|\tilde x\|_{\rm Li}$. If $\tilde x_1,\,\tilde x_2\in C([t^*,t^*+c],\mathcal K)$ and $\|S(\tilde x_1)-S(\tilde x_2)\|_{\rm Li}=\left|\int_{t^*}^t(f(s/\varepsilon,s,\tilde x_1(s))-f(s/\varepsilon,s,\tilde x_2(s)))\,ds\right|e^{-(t-t^*)L_{\mathcal K}}$ for a $t\in[t^*,t^*+c]$,

$$\begin{aligned}\|S(\tilde x_1)-S(\tilde x_2)\|_{\rm Li}&\leq e^{-(t-t^*)L_{\mathcal K}}\int_{t^*}^t L_{\mathcal K}|\tilde x_1(s)-\tilde x_2(s)|\,ds\\ &\leq e^{-(t-t^*)L_{\mathcal K}}\|\tilde x_1-\tilde x_2\|_{\rm Li}\int_{t^*}^t L_{\mathcal K}\,e^{(s-t^*)L_{\mathcal K}}ds\leq(1-e^{-cL_{\mathcal K}})\|\tilde x_1-\tilde x_2\|_{\rm Li}\,,\end{aligned}$$

which completes the proof. □

The next theorem extends to the more general setting here considered the information of [7, Theorem 3.1].

Theorem 2.8. *Assume Hypotheses 2.1 as well as $|f| \le M$ on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^n$. Fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact, and define the compact set $\tilde{\mathcal{K}} := \{x\in\mathbb{R}^n \mid \mathrm{d}(x,\mathcal{K}) \le cM\}$. Let $M_{\tilde{\mathcal{K}}}$ and $L_{\tilde{\mathcal{K}}}$ be the constants of Hypotheses 2.1. Let $(\hat{\Delta}(\varepsilon),\hat{\eta}(\varepsilon))$ be a uniform rate of averaging of f on $[0,c]\times\tilde{\mathcal{K}}$. Then,*

$$|x_\varepsilon(\tau^*+\tau;\tau^*,x^*) - z_\varepsilon(\tau^*+\tau;\tau^*,x^*)| \le e^{2cL_{\tilde{\mathcal{K}}}}\big((2+cL_{\tilde{\mathcal{K}}})M_{\tilde{\mathcal{K}}}\hat{\Delta}(\varepsilon) + c\,\hat{\eta}(\varepsilon)\big)$$

for every $(\tau^*,x^*)\in\mathbb{R}_+\times\mathcal{K}$, $\varepsilon>0$ and $\tau\in[0,c/\varepsilon]$. In particular, if

$$\hat{D}_{c,\mathcal{K}}(\varepsilon) := \sup_{(\tau^*,x^*)\in\mathbb{R}_+\times\mathcal{K},\,\tau\in[0,c/\varepsilon]} |x_\varepsilon(\tau^*+\tau;\tau^*,x^*) - z_\varepsilon(\tau^*+\tau;\tau^*,x^*)|,$$

then $\lim_{\varepsilon\to0^+}\hat{D}_{c,\mathcal{K}}(\varepsilon)=0$.

Proof. The proof is based on that of [7, Theorem 3.1]. We fix $\varepsilon>0$ and note that $x_\varepsilon(\tau^*+\tau;\tau^*,x^*) = \tilde{x}_\varepsilon(\varepsilon\tau^*+\varepsilon\tau;\varepsilon\tau^*,x^*)$ and $z_\varepsilon(\tau^*+\tau;\tau^*,x^*) = \tilde{z}(\varepsilon\tau^*+\varepsilon\tau;\varepsilon\tau^*,x^*)$ exist and belong to $\tilde{\mathcal{K}}$ for every $(\tau^*,x^*)\in\mathbb{R}_+\times\mathcal{K}$ and $\tau\in[0,c/\varepsilon]$, since $|f|\le M$ and $|\hat{f}|\le M$ on $\mathbb{R}_+\times\mathbb{R}^n$. We fix $(\tau^*,x^*)\in\mathbb{R}_+\times\mathcal{K}>0$, call $t^* := \varepsilon\,\tau^*$, call $\tilde{x}_\varepsilon(t) := \tilde{x}_\varepsilon(t;t^*,x^*)$ and $\tilde{z}(t) := \tilde{z}(t;t^*,x^*)$, define the maps S and $\hat{S}$ on $C([t^*,t^*+c],\tilde{\mathcal{K}})$ by (2.6), take $t\in[t^*,\,t^*+c]$ with $\|S(\tilde{z})-\hat{S}(\tilde{z})\|_\infty = |S(\tilde{z})(t)-\hat{S}(\tilde{z})(t)|$, and divide the interval $[t^*,\,t]$ in the subintervals $\mathcal{I}_0 := [t^*,\,t^*+\hat{\Delta}(\varepsilon))$, $\mathcal{I}_1 := [t^*+\hat{\Delta}(\varepsilon),\,t^*+2\,\hat{\Delta}(\varepsilon))$, …, $\mathcal{I}_{m_t} := [t^*+m_t\,\hat{\Delta}(\varepsilon),\,t\,]$ for $m_t\ge0$, where the length of $\mathcal{I}_{m_t}$ is at most $\hat{\Delta}(\varepsilon)$. Let t_j be the middle point of $\mathcal{I}_j$. Since all the intervals are of length at most c, Remark 2.2 and the change of variable $s=\varepsilon\,\sigma$ in (2.5) yields

$$\begin{aligned}
\|S(\tilde{z})-\hat{S}(\tilde{z})\|_\infty &= \left|\int_{t_*}^{t}(f(s/\varepsilon,s,\tilde{z}(s)) - \hat{f}(s,\tilde{z}(s)))\,ds\right| \le \sum_{j=0}^{m_t}\int_{\mathcal{I}_j}|f(s/\varepsilon,s,\tilde{z}(s)) - f(s/\varepsilon,s,\tilde{z}(t_j))|\,ds\\
&\quad+\sum_{j=0}^{m_t}\int_{\mathcal{I}_j}|\hat{f}(s,\tilde{z}(t_j)) - \hat{f}(s,\tilde{z}(s))|\,ds + \sum_{j=0}^{m_t}\left|\int_{\mathcal{I}_j}(f(s/\varepsilon,s,\tilde{z}(t_j)) - \hat{f}(s,\tilde{z}(t_j)))\,ds\right|\\
&\le 2\sum_{j=0}^{m_t}\int_{\mathcal{I}_j}L_{\tilde{\mathcal{K}}}\,|\tilde{z}(s)-\tilde{z}(t_j)|\,ds + \sum_{j=0}^{m_t-1}\hat{\eta}(\varepsilon)\hat{\Delta}(\varepsilon) + 2M_{\tilde{\mathcal{K}}}\hat{\Delta}(\varepsilon)\\
&\le 2\sum_{j=0}^{m_t}\int_{\mathcal{I}_j}L_{\tilde{\mathcal{K}}}M_{\tilde{\mathcal{K}}}|s-t_j|\,ds + c\,\hat{\eta}(\varepsilon) + 2M_{\tilde{\mathcal{K}}}\hat{\Delta}(\varepsilon) \le c\,L_{\tilde{\mathcal{K}}}M_{\tilde{\mathcal{K}}}\hat{\Delta}(\varepsilon) + c\,\hat{\eta}(\varepsilon) + 2M_{\tilde{\mathcal{K}}}\hat{\Delta}(\varepsilon)\,.
\end{aligned}$$

So, the first assertion of Lemma 2.7 ensures that $\|S(\tilde{z})-\hat{S}(\tilde{z})\|_{\mathrm{Li}} \le \alpha := (2+cL_{\tilde{\mathcal{K}}})M_{\tilde{\mathcal{K}}}\hat{\Delta}(\varepsilon)+c\,\hat{\eta}(\varepsilon)$. Since $\tilde{x}_\varepsilon = S(\tilde{x}_\varepsilon)$ and $\tilde{z}=\hat{S}(\tilde{z})$, the last assertion in Lemma 2.7 yields $\|\tilde{x}_\varepsilon-\tilde{z}\|_{\mathrm{Li}} \le \|S(\tilde{x}_\varepsilon)-S(\tilde{z})\|_{\mathrm{Li}} + \|S(\tilde{z})-\hat{S}(\tilde{z})\|_{\mathrm{Li}} \le (1-e^{-cL_{\tilde{\mathcal{K}}}})\|\tilde{x}_\varepsilon-\tilde{z}\|_{\mathrm{Li}}+\alpha$, so that $\|\tilde{x}_\varepsilon-\tilde{z}\|_{\mathrm{Li}} \le e^{cL_{\tilde{\mathcal{K}}}}\alpha$. Using again Lemma 2.7 and writing the value of α, we get $\|\tilde{x}_\varepsilon-\tilde{z}\|_\infty \le e^{2cL_{\tilde{\mathcal{K}}}}\big((2+cL_{\tilde{\mathcal{K}}})M_{\tilde{\mathcal{K}}}\hat{\Delta}(\varepsilon)+c\,\hat{\eta}(\varepsilon)\big)$. This bound is irrespective of the choices of t^* (i.e., of τ^*) and $x^*\in\mathcal{K}$, and so (2.3) proves the first assertion. Since Proposition 2.6 ensures the existence of a uniform rate of averaging of f on $[0,c]\times\tilde{\mathcal{K}}$, the second assertion follows from the first one. □

An easy adaptation of the previous proof leads to the next result, which provides a better bound for the distance between two specific solutions taking values on a previously fixed compact set:

Theorem 2.9. *Assume Hypotheses 2.1, and fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact. Let $M_{\mathcal{K}}$ and $L_{\mathcal{K}}$ be the constants of Hypotheses 2.1. Let $(\hat{\Delta}(\varepsilon),\hat{\eta}(\varepsilon))$ be a uniform rate of averaging of f on $[0,c]\times\mathcal{K}$. If $\varepsilon>0$, $\tau^*\ge0$, and $x_\varepsilon(\tau^*+\tau;\tau^*,x^*)$ and $z_\varepsilon(\tau^*+\tau;\tau^*,x^*)$ exist and belong to $\mathcal{K}$ for every $\tau\in[0,c/\varepsilon]$, then*

$$|x_\varepsilon(\tau^*+\tau;\tau^*,x^*) - z_\varepsilon(\tau^*+\tau;\tau^*,x^*)| \le e^{2cL_{\mathcal{K}}}((2+cL_{\mathcal{K}})M_{\mathcal{K}}\hat{\Delta}(\varepsilon) + c\,\hat{\eta}(\varepsilon))$$

for every $\tau\in[0,c/\varepsilon]$.

2.2. Nonuniform averaging results

Less exigent conditions give rise to weaker results on averaging.

Hypotheses 2.10. The map $f:\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^n\to\mathbb{R}^n$ is continuous and, for every $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact: there exists $M_{c,\mathcal{K}}\ge0$ such that $|f|\le M_{c,\mathcal{K}}$ on $\mathbb{R}_+\times[0,c]\times\mathcal{K}$; there exist a modulus of continuity $\theta_{c,\mathcal{K}}$ and $L_{c,\mathcal{K}}>0$ such that $|f(\tau,t_1,x_1)-f(\tau,t_2,x_2)| \le \theta_{c,\mathcal{K}}(|t_1-t_2|)+L_{c,\mathcal{K}}\,|x_1-x_2|$ for all $\tau\in\mathbb{R}_+$, $t_1,t_2\in[0,c]$ and $x_1,x_2\in\mathcal{K}$; and there exists $\hat{f}(t,x) := \lim_{T\to\infty}(1/T)\int_0^T f(s,t,x)\,ds$ for all $(t,x)\in\mathbb{R}\times\mathbb{R}^n$.

Remark 2.11. It is easy to deduce that $|\hat{f}(t_1,x_1)| \le M_{c,\mathcal{K}}$ and $|\hat{f}(t_1,x_1)-\hat{f}(t_2,x_2)| \le \theta_{c,\mathcal{K}}(|t_1-t_2|)+L_{c,\mathcal{K}}\,|x_1-x_2|$ for all $t_1,t_2\in[0,c]$ and $x_1,x_2\in\mathcal{K}$. In particular, $\hat{f}$ is uniformly continuous on $[0,c]\times\mathcal{K}$. In addition, if $\hat{f}_T(t,x) = (1/T)\int_0^T f(\sigma,t,x)\,d\sigma$, then $\hat{f}(t,x) = \lim_{T\to\infty}\hat{f}_T(t,x)$ uniformly on $[0,c]\times\mathcal{K}$: it is easy to check that the family $\{\hat{f}_T \mid T\in(0,\infty)\}\subset C([0,c]\times\mathcal{K},\mathbb{R}^n)$ is uniformly bounded by $M_{c,\mathcal{K}}$ and equicontinuous, and hence the assertion follows from Arzelà-Ascoli's Theorem.

A map satisfying Hypotheses 2.10 is often called a *KBM* map (after Krylov, Bogoliubov and Mitropolsky): see [6]. Analogously, we call *UKBM* maps (with U of uniform) to those maps satisfying Hypotheses 2.1.

Definition 2.12. Assume Hypotheses 2.10, and fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact. The *static rate of convergence of f on* $[0,c]\times\mathcal{K}$ is

$$\delta^S_{c,\mathcal{K}}(\varepsilon) := \sup_{(t,x)\in[0,c]\times\mathcal{K},\,T\in[0,c/\varepsilon]} \varepsilon\left|\int_0^T (f(\sigma,t,x)-\hat f(t,x))\,d\sigma\right|,$$

and the *dynamic rate of convergence of f on* $[0,c]\times\mathcal{K}$ is

$$\delta^D_{c,\mathcal{K}}(\varepsilon) := \sup_{x\in\mathcal{K},\,T\in[0,c/\varepsilon]} \varepsilon\left|\int_0^T (f(\sigma,\varepsilon\sigma,x)-\hat f(\varepsilon\sigma,x))\,d\sigma\right|.$$

Proposition 2.13. *Assume Hypotheses 2.10, and fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact. Then,*

(i) $\lim_{\varepsilon\to0^+}\delta^S_{c,\mathcal{K}}(\varepsilon)=0$, (ii) $\lim_{\varepsilon\to0^+}\delta^D_{c,\mathcal{K}}(\varepsilon)=0$.

Proof. (i) The proof repeats in this simpler setting that of Proposition 2.4(i), having in mind the uniform convergence established in Remark 2.11: we work for $\tau^*=0$ fixed instead of letting τ^* vary in $\mathbb{R}_+$, and with the bound $M_{c,\mathcal{K}}$ of $|f|$ on $\mathbb{R}_+\times[0,c]\times\mathcal{K}$.

(ii) We call $\delta_\varepsilon := \delta^S_{c,\mathcal{K}}(\varepsilon)$. Note that $\delta_\varepsilon = \sup_{(t,x)\in[0,c]\times\mathcal{K},T\in[0,c]}\left|\int_0^T (f(s/\varepsilon,t,x)-\hat f(t,x))\,ds\right|$. So, if $0\le\bar t\le\bar t+\sqrt{\delta_\varepsilon}\le c$, then $\left|\int_{\bar t}^{\bar t+\sqrt{\delta_\varepsilon}}(f(s/\varepsilon,t,x)-\hat f(t,x))\,ds\right|\le 2\,\delta_\varepsilon$ for every $(t,x)\in[0,c]\times\mathcal{K}$, since $|\int_{\bar t}^{\bar t+\sqrt{\delta_\varepsilon}}|\le|\int_0^{\bar t}|+|\int_0^{\bar t+\sqrt{\delta_\varepsilon}}|$. This bound and $M_{c,\mathcal{K}}$ substitute (2.4) and $M_\mathcal{K}$ in the adaptation to this setting of the proof of Proposition 2.4(ii): we work with $t_0=0$ instead of letting t_0 vary in $\mathbb{R}_+$, and obtain $\left|\int_0^T(f(s/\varepsilon,s,x)-\hat f(s,x))\,ds\right|\le 2\,c\,\theta_{c,\mathcal{K}}(\sqrt{\delta_\varepsilon})+(2c+2M_{c,\mathcal{K}})\sqrt{\delta_\varepsilon}$ for all $x\in\mathcal{K}$ and $T\in[0,c]$, where $\lim_{\delta\to0^+}\theta_{c,\mathcal{K}}(\delta)=0$. The results follows from here and (i), since $\delta^D_{c,\mathcal{K}}(\varepsilon)=\sup_{x\in\mathcal{K},T\in[0,c]}\left|\int_0^T(f(s/\varepsilon,s,x)-\hat f(s,x))\,ds\right|$. □

Definition 2.14. Assume Hypotheses 2.10, and fix $c>0$ and $\mathcal{X}\subseteq\mathbb{R}^n$. A pair of nonnegative functions $(\Delta(\varepsilon),\eta(\varepsilon))$ is a *rate of averaging of f on* $[0,c]\times\mathcal{X}$ if $\lim_{\varepsilon\to0^+}\Delta(\varepsilon)=\lim_{\varepsilon\to0^+}\eta(\varepsilon)=0$ and

$$\frac{\varepsilon}{\Delta(\varepsilon)}\left|\int_\tau^{\tau+\Delta(\varepsilon)/\varepsilon}(f(\sigma,\varepsilon\sigma,x)-\hat f(\varepsilon\sigma,x))\,d\sigma\right|\le\eta(\varepsilon) \tag{2.7}$$

for every $(\tau,x)\in\mathbb{R}_+\times\mathcal{X}$ and every $\varepsilon\in(0,\infty)$ with $\varepsilon\tau+\Delta(\varepsilon)\in[0,c]$.

Proposition 2.15. *Assume Hypotheses 2.10, and fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact. For any $\alpha\in(0,1)$, the pair $\big((\delta^D_{c,\mathcal{K}}(\varepsilon))^\alpha,\,2\,(\delta^D_{c,\mathcal{K}}(\varepsilon))^{1-\alpha}\big)$ is a rate of averaging of f on $[0,c]\times\mathcal{K}$.*

Proof. Proposition 2.13(ii) ensures that both maps tend to 0 as $\varepsilon\to0^+$. Since $|\int_\tau^{\tau+\delta/\varepsilon}|\le|\int_0^\tau|+|\int_0^{\tau+\delta/\varepsilon}|$,

$$\varepsilon\left|\int_\tau^{\tau+(\delta^D_{c,\mathcal{K}}(\varepsilon))^\alpha/\varepsilon}(f(\sigma,\varepsilon\sigma,x)-\hat f(\varepsilon\sigma,x))\,d\sigma\right|\le 2\,\delta^D_{c,\mathcal{K}}(\varepsilon)$$

for every $(\tau,x)\in\mathbb{R}_+\times\mathcal{K}$ if $0\le\tau+(\delta^D_{c,\mathcal{K}}(\varepsilon))^\alpha/\varepsilon\le c/\varepsilon$. So, we divide both sides by $(\delta^D_{c,\mathcal{K}}(\varepsilon))^\alpha$. □

Theorem 2.16. *Assume Hypotheses 2.10 as well as $|f|\le M$ on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^n$. Fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact, and define the compact set $\tilde{\mathcal{K}} := \{x\in\mathbb{R}^n \mid \mathrm{d}(x,\mathcal{K})\le cM\}$. Let $M_{c,\tilde{\mathcal{K}}}$ and $L_{c,\tilde{\mathcal{K}}}$ be the constants of Hypotheses 2.10. Let $(\Delta(\varepsilon),\eta(\varepsilon))$ be a rate of averaging of f on $[0,c]\times\tilde{\mathcal{K}}$. Then,*

$$|x_\varepsilon(\tau^*+\tau;\tau^*,x^*)-z_\varepsilon(\tau^*+\tau;\tau^*,x^*)|\le e^{2cL_{c,\tilde{\mathcal{K}}}}((2+c\,L_{c,\tilde{\mathcal{K}}})M_{c,\tilde{\mathcal{K}}}\Delta(\varepsilon)+c\,\eta(\varepsilon))$$

for every $(\tau^*,x^*)\in\mathbb{R}_+\times\mathcal{K}$, $\varepsilon>0$ and $\tau\in[0,c/\varepsilon]$. In particular, if

$$D_{c,\mathcal{K}}(\varepsilon) := \sup_{x^*\in\mathcal{K},\,\tau\in[0,c/\varepsilon]}|x_\varepsilon(\tau;0,x^*)-z_\varepsilon(\tau;0,x^*)|,$$

then $\lim_{\varepsilon\to0^+}D_{c,\mathcal{K}}(\varepsilon)=0$.

Proof. Once again, we omit the details of the proof, which basically uses the results of this section to repeat the proof of Theorem 2.8 (slightly more complicated than this one) with $t^*=0$ and with (Δ,η) instead of $(\hat\Delta,\hat\eta)$. □

Finally, as in the uniform case, the previous proof can be easily adapted to check the next result.

Theorem 2.17. *Assume Hypotheses 2.10, and fix $c>0$ and $\mathcal{K}\subset\mathbb{R}^n$ compact. Let $M_{c,\mathcal{K}}$ and $L_{c,\mathcal{K}}$ be the constants of Hypotheses 2.10. Let $(\Delta(\varepsilon),\eta(\varepsilon))$ be a rate of averaging of f on $[0,c]\times\mathcal{K}$. If $\varepsilon>0$ and $x_\varepsilon(\tau;0,x^*)$ and $z_\varepsilon(\tau;0,x^*)$ exist and belong to $\mathcal{K}$ for every $\tau\in[0,c/\varepsilon]$, then*

$$\left|x_\varepsilon(\tau;0,x^*)-z_\varepsilon(\tau;0,x^*)\right|\le e^{2cL_{c,\mathcal{K}}}((2+c\,L_{c,\mathcal{K}})M_{c,\mathcal{K}}\Delta(\varepsilon)+c\,\eta(\varepsilon))$$

for every $\tau\in[0,c/\varepsilon]$.

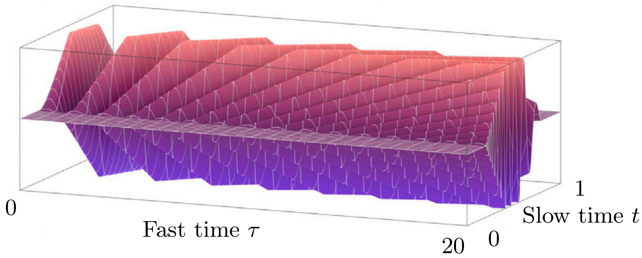

Fig. 1. The map $(\tau, t) \mapsto \phi(t)\sin(\tau/t)$ of Section 2.3.

2.3. Optimality of the hypotheses

Let ϕ be the continuous piecewise linear function which takes value 1 in $[2/5, 3/5]$ and value 0 outside $[1/5, 4/5]$. It is easy to check that the globally bounded map $f(\tau, t, x) := \phi(t)\sin(\tau/t)$ (with $f(\tau, 0, x) = 0$) satisfies all the conditions of Hypotheses 2.10 (with mean function $\hat{f} \equiv 0$) excepting the equicontinuity on $[0, c]$ of $\{t \mapsto \phi(t)\sin(\tau/t) \mid \tau \in \mathbb{R}_+\}$ for $c > 1/5$. The depiction of the graph of f in Fig. 1 shows the reason for this: the map $t \mapsto f(\tau, t)$ oscillates faster and faster as τ increases. It is easy to check that the dynamic rate of convergence (which is independent of x, as f) is given by $\delta_c^D(\varepsilon) = \sup_{T\in[0,c]} \left| \int_0^T \phi(s)\sin(1/\varepsilon)\, ds \right| = |\sin(1/\varepsilon)| \int_0^c \phi(s)\, ds$, which does not tend to 0 as $\varepsilon \to 0^+$ if $c > 1/5$. In addition, $\sup_{\tau\in[0,c/\varepsilon]} \left| x_\varepsilon(\tau; 0, x^*) - z_\varepsilon(\tau; 0, x^*) \right| = |\sin(1/\varepsilon)| \int_0^c \phi(s)\, ds$, and hence the limit is not 0 whenever $c > 1/5$. So, the last assertion of Theorem 2.16 does not hold under these weaker conditions.

Data availability

No data was used for the research described in the article.

References

[1] N. Bogoliubov, Y.A. Mitropolsky, Asymptotic Methods in the Theory of Non-linear Oscillations, Gordon & Breach, 1961, (english translation).
[2] J. Hale, Ordinary Differential Equations, Wiley-Interscience, 1969.
[3] V. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations, second ed., Springer, 1988.
[4] Y.A. Mitropolsky, Problems of the Asymptotic Theory of Nonstationary Vibrations, 1965, Israel Program for Scientific Translations.
[5] J. Sanders, F. Verhulst, Averaging Methods in Nonlinear Dynamical Systems, Springer, 1985.
[6] J. Sanders, F. Verhulst, J. Murdock, Averaging Methods in Nonlinear Dynamical Systems, Springer Nature, 2007.
[7] Z. Artstein, Averaging of time-varying differential equations revisited, J. Differential Equations 243 (2) (2007) 146–167.
[8] I. Bright, Tight estimates for general averaging applied to almost-periodic differential equations, J. Differential Equations 246 (2009) 2922—2937.
[9] Z. Artstein, Averaging of ordinary differential equations with slowly varying averages, 14, (2) 2010, pp. 353–365,
[10] I. Longo, R. Obaya, A. Sanz, Nonautonomous modelling in energy balance models of climate. Limitations of averaging and climate sensitivity, 485, 2026, 135038,