Averaging dynamic equations on time scales*

Antonín Slavík
Charles University, Faculty of Mathematics and Physics,
Sokolovská 83, 186 75 Praha 8, Czech Republic
E-mail: slavik@karlin.mff.cuni.cz

Abstract

The aim of this paper is to generalize the classical theorems on averaging of differential equations. We focus on dynamic equations on time scales and prove both periodic and nonperiodic version of the averaging theorem, as well as a related theorem on the existence of periodic solutions.

Keywords: Averaging method, periodic averaging, periodic time scales, periodic solutions, nonperiodic averaging, generalized ordinary differential equations

MSC 2010 subject classification: 34C29, 26E70, 34N05, 34C25, 39A12, 39A23

1 Introduction

The averaging method for ordinary differential equations has its roots in the works of Lagrange and Laplace on celestial mechanics. There are many sources which describe the method and its applications; a modern and detailed treatment is given in the specialized monograph [15] or in its second revised edition [16]. The book [17] contains a chapter devoted to averaging, which can serve as a first introduction with many examples. The simplest version of averaging is concerned with the initial-value problem

$$x'(t) = \varepsilon f(t, x(t)), \quad x(t_0) = x_0,$$

where $\varepsilon > 0$ is a small parameter. The basic idea is that because x is a slowly varying function, we can obtain an approximate solution by averaging the right-hand side f with respect to t while holding x constant, i.e. we consider the equation

$$y'(t) = \varepsilon f^0(y(t)), \quad y(t_0) = x_0,$$

where

$$f^0(y) = \frac{1}{T} \int_{t_0}^{t_0+T} f(t, y)\, \mathrm{d}t$$

if f is T-periodic in the first argument and

$$f^0(y) = \lim_{T\to\infty} \frac{1}{T} \int_{t_0}^{t_0+T} f(t, y)\, \mathrm{d}t$$

*Supported by grant KJB101120802 of the Grant Agency of the Academy of Sciences of the Czech Republic, and by grant MSM 0021620839 of the Czech Ministry of Education.

otherwise. According to the classical averaging theorems, the solution of the averaged equation (which is autonomous and therefore easier to analyze) provides a good approximation to the original solution for $t \in [t_0, t_0 + d/\varepsilon]$, where d is a certain constant.

Our aim is to obtain both periodic and nonperiodic version of the averaging theorem for dynamic equations on time scales. Note that our term "time scale" has a different meaning than in [15] and [16] and denotes a closed subset of the real line. We assume that the reader is familiar with the basic notions of calculus on time scales as presented in [4] and with integration theory on time scales as presented in [5]. Our proof of the periodic averaging theorem follows the proof of the classical theorem from [16]. We also prove a related theorem on the existence of periodic solutions, which generalizes another well-known result from the theory of ordinary differential equations. The proof is also inspired by the classical one (see e.g. [17]), but certain technical details are more complicated.

Finally, we obtain a nonperiodic version of the averaging theorem by converting the dynamic equations to generalized ordinary differential equations and using an existing theorem on averaging of generalized equations. In this part, some familiarity with the theory of generalized ordinary differential equations might be helpful (see e.g. the book [12] and the paper [13], which describes the correspondence between dynamic equations and generalized ordinary differential equations).

Our averaging theorems unify and extend existing results for differential and difference equations (periodic and nonperiodic averaging theorems for difference equations can be found in [6] and [1], respectively). As far as the author knows, the theorem on the existence of periodic solutions is new even in the purely discrete case.

2 Auxiliary results

Let $\mathbb{T}$ be a time scale, i.e. a nonempty closed subset of $\mathbb{R}$. For every $t \in \mathbb{T}$, we define the forward jump operator by $\sigma(t) = \inf\{s \in \mathbb{T},\ s > t\}$ and the graininess function by $\mu(t) = \sigma(t) - t$. If $\sigma(t) > t$, we say that t is right-scattered; otherwise, t is right-dense. Similarly, the backward jump operator is given by $\rho(t) = \sup\{s \in \mathbb{T},\ s < t\}$, and we distinguish between left-scattered and left-dense points depending on whether $\rho(t) < t$ or $\rho(t) = t$. If $\mathbb{T}$ has a left-scattered maximum M, then we define $\mathbb{T}^\kappa = \mathbb{T} - \{M\}$; otherwise, $\mathbb{T}^\kappa = \mathbb{T}$.

Given a pair of numbers $a, b \in \mathbb{T}$, the symbol $[a, b]_{\mathbb{T}}$ will be used to denote a closed interval in $\mathbb{T}$, i.e. $[a, b]_{\mathbb{T}} = \{t \in \mathbb{T};\ a \le t \le b\}$. On the other hand, $[a, b]$ is the usual closed interval on the real line, i.e. $[a, b] = \{t \in \mathbb{R};\ a \le t \le b\}$. This notational convention should help the reader to distinguish between ordinary and time scale intervals.

A function $f : \mathbb{T} \to \mathbb{R}$ is called rd-continuous if it is regulated on $\mathbb{T}$ and continuous at right-dense points of $\mathbb{T}$.

In the time scale calculus, the usual derivative $f'(t)$ is replaced by the Δ-derivative $f^\Delta(t)$, where $t \in \mathbb{T}^\kappa$. Similarly, the usual integral $\int_a^b f(t)\,\mathrm{d}t$ is replaced by the Δ-integral $\int_a^b f(t)\,\Delta t$, where $f : [a, b]_{\mathbb{T}} \to \mathbb{R}$. The definitions and properties of the Δ-derivative and Δ-integral can be found in [4] and [5]. We remark that the notion of Riemann Δ-integral is sufficient for our purposes, although time scale versions of the more general Lebesgue and Kurzweil integrals are available, too.

Definition 2.1. Let $T > 0$ be a real number. A time scale $\mathbb{T}$ is called T-periodic if $t \in \mathbb{T}$ implies $t+T \in \mathbb{T}$ and $\mu(t) = \mu(t + T)$.

For example, the time scale $\mathbb{T} = \bigcup_{k=0}^\infty [2k, 2k + 1]$ is 2-periodic, but also 4-periodic etc.

The condition $\mu(t) = \mu(t+T)$ is important; as the following lemma shows, it guarantees that if $t \in \mathbb{T}$, then the interval $[t + T, t + 2T]_{\mathbb{T}}$ does not contain more points than $[t, t + T]_{\mathbb{T}}$.

Lemma 2.2. *Let $\mathbb{T}$ be a T-periodic time scale. If $t \in \mathbb{T}$, then the function $s(\tau) = \tau + T$ maps $[t, t + T]_{\mathbb{T}}$ onto $[t + T, t + 2T]_{\mathbb{T}}$.*

Proof. Let $O = [t, t+T] \backslash \mathbb{T}$. Since O is an open set, it can be written as a countable union of disjoint open intervals, i.e. $O = \bigcup_n (a_n, b_n)$. Assume there exists a $\tau \in [t+T, t+2T]_{\mathbb{T}}$ which is not contained in the range of s, i.e. $\tau - T \in O$. It follows that $\tau - T \in (a_n, b_n)$ for a certain n; note that $a_n, b_n \in \mathbb{T}$ and $\mu(a_n) = b_n - a_n$. But $\tau \in (a_n + T, b_n + T)$, and consequently $\mu(a_n + T) \leq \tau - a_n < b_n - a_n$, a contradiction. □

Lemma 2.3. *Let $\mathbb{T}$ be a T-periodic time scale and $f : \mathbb{T} \to \mathbb{R}$ a T-periodic function. If $a, b \in \mathbb{T}$, $a \leq b$, and f is Δ-integrable on $[a, b]_{\mathbb{T}}$, then for every $k \in \mathbb{N}$, f is Δ-integrable on $[a + kT, b + kT]_{\mathbb{T}}$ and*

$$\int_a^b f(\tau)\Delta\tau = \int_{a+kT}^{b+kT} f(\tau)\Delta\tau.$$

Proof. The statement follows immediately from the definition of the Δ-integral and the fact that, according to Lemma 2.2, there is a one-to-one correspondence between the partitions of $[a, b]_{\mathbb{T}}$ and the partitions of $[a + kT, b + kT]_{\mathbb{T}}$. □

It follows from the previous lemma and the additivity of the Δ-integral that integrating a T-periodic function f over an arbitrary interval of length T always gives the same result.

Let an arbitrary time scale $\mathbb{T}$ be given. If $t_0 \in \mathbb{T}$ and $f : \mathbb{T} \to \mathbb{R}$ is a regressive function (i.e., $1 + \mu(t)f(t) \neq 0$ for every t), it is known that the initial value problem

$$y^{\Delta}(t) = f(t)y(t), \quad y(t_0) = 1$$

has a unique solution on $\mathbb{T}$. This solution is called the exponential function corresponding to t_0 and f and its value at $t \in \mathbb{T}$ will be denoted by $e_f(t, t_0)$. (For more details, see [4]. A more general definition of the exponential function applicable to Riemann Δ-integrable matrix functions f is given in [14].)

The following theorem from [4, Corollary 6.8] represents a Gronwall-type inequality.

Theorem 2.4. *Let $y : \mathbb{T} \to \mathbb{R}$ be a rd-continuous function and $t_0 \in \mathbb{T}$. Suppose there exist constants α, β, $\gamma \in \mathbb{R}$, $\gamma > 0$, such that*

$$y(t) \leq \alpha + \beta(t - t_0) + \gamma \int_{t_0}^{t} y(s)\,\Delta s$$

for every $t \in \mathbb{T}$, $t \geq t_0$. Then

$$y(t) \leq (\alpha + \beta/\gamma)\, e_{\gamma}(t, t_0) - \beta/\gamma$$

for every $t \in \mathbb{T}$, $t \geq t_0$.

The following theorem gives an estimate for the exponential function.

Theorem 2.5. *Let $f : \mathbb{T} \to \mathbb{R}$ be a regressive rd-continuous function such that $|f(t)| \leq C$ for every $t \in \mathbb{T}$. Then $|e_f(t, t_0)| \leq e^{C(t-t_0)}$ for each pair of values t, $t_0 \in \mathbb{T}$, $t \geq t_0$.*

Proof. According to [14], the value of $e_f(t, t_0)$ is equal to the product integral of f over $[t_0, t]_{\mathbb{T}}$. This means that for every $\varepsilon > 0$, there is a partition $D : t_0 = s_0 < s_1 < \cdots < s_m = t$ of interval $[t_0, t]_{\mathbb{T}}$ (where $s_i \in \mathbb{T}$ for every i) such that the product

$$P(f, D) = \prod_{i=1}^{m} (1 + f(t_{i-1})(t_i - t_{i-1}))$$

satisfies $|e_f(t, t_0) - P(f, D)| < \varepsilon$. Thus

$$|e_f(t, t_0)| \leq |e_f(t, t_0) - P(f, D)| + |P(f, D)| < \varepsilon + \prod_{i=1}^{m} (1 + C(t_i - t_{i-1}))$$

$$\leq \varepsilon + \prod_{i=1}^{m} e^{C(t_i - t_{i-1})} = \varepsilon + e^{C(t-t_0)}.$$

The statement follows from the fact that ε can be arbitrarily small. □

In fact, we will need the last theorem only in the case when f is a constant function. However, the statement as well as the proof can be easily generalized for Riemann Δ-integrable matrix functions $f : \mathbb{T} \to \mathbb{R}^{n\times n}$.

Consider a function $f : \mathbb{T} \times \mathbb{R}^n \to \mathbb{R}^n$. For a fixed $x \in \mathbb{R}^n$, the function $t \mapsto f(t, x)$ is defined on $\mathbb{T}$ and its Δ-derivative (provided it exists) will be denoted by f^Δ. On the other hand, for a fixed $t \in \mathbb{T}$, the function $x \mapsto f(t, x)$ is defined on $\mathbb{R}^n$ and its partial derivatives (provided they exist) will be denoted by $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$; the symbol $\frac{\partial f}{\partial x}$ will be used for the differential (i.e., the Jacobian matrix) of f.

We need the following theorem concerning differentiation under the integral sign with respect to a parameter; we omit the proof since it is identical to the classical one, and since the time scale version (in a slightly less general form) was already proved in [2].

Theorem 2.6. *If $U \subset \mathbb{R}^n$ is an open set and $f : [a, b]_\mathbb{T} \times U \to \mathbb{R}$ is a continuous function with continuous partial derivatives with respect to $x_1, \dots, x_n$, then the function*

$$g(x) = \int_a^b f(t, x)\,\Delta t, \quad x \in U,$$

is continuous and has continuous partial derivatives in U given by the formula

$$\frac{\partial g}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(t, x)\,\Delta t, \;\; i \in \{1, \dots, n\}.$$

It is easy to see that the previous statement is true even for vector-valued functions $f : [a, b]_\mathbb{T} \times U \to \mathbb{R}^n$.

The following mean-value theorem is proved in [5] (see Theorem 1.14).

Theorem 2.7. *Let $f : [a, b]_\mathbb{T} \to \mathbb{R}$ be a continuous function, which is Δ-differentiable on $[a, b)_\mathbb{T}$. Then there exist numbers ξ, $\tau \in [a, b)_\mathbb{T}$ such that*

$$f^\Delta(\tau) \leq \frac{f(b) - f(a)}{b - a} \leq f^\Delta(\xi).$$

We need the following statement, which is an easy corollary of the previous theorem.

Corollary 2.8. *Let $f : [a, b]_\mathbb{T} \to \mathbb{R}^n$ be a continuous function, which is Δ-differentiable on $[a, b)_\mathbb{T}$. If there exists a number $M \geq 0$ such that $\|f^\Delta(t)\| \leq M$ for every $\tau \in [a, b)_\mathbb{T}$, then $\|f(b) - f(a)\| \leq M(b-a)$.*

A time scale version of the implicit function theorem will be used in section 4. The following result where the implicit function is described by a single equation is proved in section 9 of [3]. The notion of a partial Δ-derivative of a multivariable time scale function is introduced in the same paper.

Theorem 2.9. *Consider time scales $\mathbb{T}_1, \dots, \mathbb{T}_l$, a point $p^0 = (t_1^0, \dots, t_l^0, y^0) \in \mathbb{T}_1 \times \dots \times \mathbb{T}_l \times \mathbb{R}$ and a function $F : U \to \mathbb{R}$, where U is a neighborhood of p^0. Assume that F satisfies the following conditions:*

1. *F and $\frac{\partial F}{\partial y}$ are continuous in U.*
2. *$F(p^0) = 0$.*
3. *$\frac{\partial F}{\partial y}(p^0) \neq 0$.*

Then there is an open set $V \subset \mathbb{T}_1 \times \cdots \times \mathbb{T}_l$ containing the point $(t_1^0, \dots, t_l^0)$ and a unique continuous function $\psi : V \to \mathbb{R}$ such that $y^0 = \psi(t_1^0, \dots, t_l^0)$ and

$$F(t_1, \dots, t_l, \psi(t_1, \dots, t_l)) = 0$$

for every $(t_1, \dots, t_l) \in V$. Moreover, existence and continuity of the partial Δ-derivatives $\frac{\partial F}{\Delta_1 t_1}, \dots, \frac{\partial F}{\Delta_l t_l}$ in $U \cap \mathbb{T}_1^\kappa \times \cdots \times \mathbb{T}_l^\kappa$ implies the existence of $\frac{\partial \psi}{\Delta_1 t_1}, \dots, \frac{\partial \psi}{\Delta_l t_l}$ in $V \cap \mathbb{T}_1^\kappa \times \cdots \times \mathbb{T}_l^\kappa$.

In the following more general result, the implicit function is given by a system of equations (i.e., a single vector equation). We omit the proof since it is the same as in the classical case; it is sufficient to use Theorem 2.9 and proceed by induction on the number of equations (see e.g. section 3.2 in [8]).

Theorem 2.10. *Consider time scales $\mathbb{T}_1, \dots, \mathbb{T}_l$, a point $p^0 = (t_1^0, \dots, t_l^0, y_1^0, \dots, y_n^0) \in \mathbb{T}_1 \times \cdots \times \mathbb{T}_l \times \mathbb{R}^n$, and a function $F : U \to \mathbb{R}^n$, where U is a neighborhood of p^0. Assume that F satisfies the following conditions:*

1. *F and $\frac{\partial F}{\partial y_1}, \dots, \frac{\partial F}{\partial y_n}$ are continuous in U.*
2. *$F(p^0) = 0$.*
3. *$\det \frac{\partial F}{\partial y}(p^0) \neq 0$.*

Then there is an open set $V \subset \mathbb{T}_1 \times \cdots \times \mathbb{T}_l$ containing the point $(t_1^0, \dots, t_l^0)$ and a unique system of continuous functions $\psi_1, \dots, \psi_n : V \to \mathbb{R}^n$ such that $y_j^0 = \psi_j(t_1^0, \dots, t_l^0)$ for every $j \in \{1, \dots, n\}$ and

$$F(t_1, \dots, t_l, \psi_1(t_1, \dots, t_l), \dots, \psi_n(t_1, \dots, t_l)) = 0$$

for every $(t_1, \dots, t_l) \in V$. Moreover, existence and continuity of the partial Δ-derivatives $\frac{\partial F}{\Delta_1 t_1}, \dots, \frac{\partial F}{\Delta_l t_l}$ in $U \cap \mathbb{T}_1^\kappa \times \cdots \times \mathbb{T}_l^\kappa$ implies the existence of $\frac{\partial \psi_i}{\Delta_j t_j}$ in $V \cap \mathbb{T}_1^\kappa \times \cdots \times \mathbb{T}_l^\kappa$ for $i \in \{1, \dots, n\}$ and $j \in \{1, \dots, l\}$.

3 Periodic averaging

Assume that $\mathbb{T}$ is a T-periodic time scale, $t_0 \in \mathbb{T}$, $U \subset \mathbb{R}^n$, and $f : [t_0, \infty)_\mathbb{T} \times U \to \mathbb{R}^n$ is a continuous function which is T-periodic in the first argument. Throughout this section, we will use the following notation:

$$f^0(x) = \frac{1}{T} \int_{t_0}^{t_0+T} f(t, x)\, \Delta t, \quad x \in U$$

We are now ready to prove the periodic averaging theorem. Our proof is very similar to the one given in [16] (see Theorem 2.8.1 and Lemma 2.8.2) for the case $\mathbb{T} = \mathbb{R}$.

Theorem 3.1. *Let $\mathbb{T}$ be a T-periodic time scale, $U \subset \mathbb{R}^n$, $t_0 \in \mathbb{T}$, $\varepsilon_0 > 0$, $d > 0$. Consider a pair of bounded continuous functions $f : [t_0, \infty)_\mathbb{T} \times U \to \mathbb{R}^n$ and $g : [t_0, \infty)_\mathbb{T} \times U \times (0, \varepsilon_0] \to \mathbb{R}^n$. Assume that f is T-periodic in the first argument and Lipschitz-continuous in the second argument. Moreover, suppose that for every $\varepsilon \in (0, \varepsilon_0]$, the initial-value problems*

$$\begin{aligned} x^\Delta(t) &= \varepsilon f(t, x(t)) + \varepsilon^2 g(t, x(t), \varepsilon), \quad x(t_0) = x_0(\varepsilon), \\ y^\Delta(t) &= \varepsilon f^0(y(t)), \quad y(t_0) = y_0(\varepsilon) \end{aligned}$$

have solutions $x_\varepsilon, y_\varepsilon : [t_0, t_0 + d/\varepsilon]_\mathbb{T} \to U$. If there is a constant $B > 0$ such that $\|x_0(\varepsilon) - y_0(\varepsilon)\| \leq B\varepsilon$ for every $\varepsilon \in (0, \varepsilon_0]$, then there exists a constant $C > 0$ such that

$$\|x_\varepsilon(t) - y_\varepsilon(t)\| \leq C\varepsilon$$

for every $\varepsilon \in (0, \varepsilon_0]$ and every $t \in [t_0, t_0 + d/\varepsilon]_\mathbb{T}$.

Proof. There exist constants L, $M > 0$ such that $\|f(t,x)\| \le M$, $\|f^0(x)\| \le M$ and $\|g(t,x,\varepsilon)\| \le M$ for every $t \in [t_0,\infty)_{\mathbb{T}}$, $x \in U$, $\varepsilon \in (0,\varepsilon_0]$, and $\|f(t,x) - f(t,y)\| \le L\|x-y\|$ whenever $t \in [t_0,\infty)_{\mathbb{T}}$, x, $y \in U$. Clearly, f^0 is Lipschitz-continuous with the same constant L. For every $\varepsilon \in (0,\varepsilon_0]$ and $t \in [t_0, t_0 + d/\varepsilon]_{\mathbb{T}}$, we have the estimate

$$\|x_\varepsilon(t) - y_\varepsilon(t)\| \le \|x_0(\varepsilon) - y_0(\varepsilon)\| + \varepsilon \int_{t_0}^{t} \|f(s, x_\varepsilon(s)) + \varepsilon g(s, x_\varepsilon(s), \varepsilon) - f^0(y_\varepsilon(s))\| \,\Delta s$$

$$\le B\varepsilon + \varepsilon \left\| \int_{t_0}^{t} (f(s, x_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s \right\| + \varepsilon^2 M(t - t_0)$$

$$\le B\varepsilon + \varepsilon \left\| \int_{t_0}^{t} (f(s, x_\varepsilon(s)) - f(s, y_\varepsilon(s)))\,\Delta s \right\| + \varepsilon \left\| \int_{t_0}^{t} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s \right\| + \varepsilon^2 M(t - t_0)$$

$$\le B\varepsilon + \varepsilon L \int_{t_0}^{t} \|x_\varepsilon(s) - y_\varepsilon(s)\|\,\Delta s + \varepsilon \left\| \int_{t_0}^{t} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s \right\| + \varepsilon^2 M(t - t_0). \quad (1)$$

Let m be the largest integer such that $t_0 + mT \le t$. Then

$$\int_{t_0}^{t} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s =$$

$$\sum_{i=1}^{m} \int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s + \int_{t_0+mT}^{t} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s. \quad (2)$$

For every $i \in \{1, \dots, m\}$, we have

$$\int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s = \int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(s)) - f(s, y_\varepsilon(t_0 + iT)))\,\Delta s$$

$$+ \int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(t_0 + iT)) - f^0(y_\varepsilon(t_0 + iT)))\,\Delta s + \int_{t_0+(i-1)T}^{t_0+iT} (f^0(y_\varepsilon(t_0 + iT)) - f^0(y_\varepsilon(s)))\,\Delta s.$$

By the definition of f^0, the second of these integrals is zero:

$$\int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(t_0 + iT)) - f^0(y_\varepsilon(t_0 + iT)))\,\Delta s = \int_{t_0}^{t_0+T} f(s, y_\varepsilon(t_0 + iT))\,\Delta s - Tf^0(y_\varepsilon(t_0 + iT)) = 0.$$

Since $y^\Delta(t) = \varepsilon f^0(y_\varepsilon(t))$ and the norm of f^0 is bounded by M, Corollary 2.8 gives

$$\|y_\varepsilon(s) - y_\varepsilon(t_0 + iT)\| \le \varepsilon M(t_0 + iT - s) \le \varepsilon MT, \quad s \in [t_0 + (i-1)T, t_0 + iT]_{\mathbb{T}}.$$

Consequently, we obtain the following estimates for the first and third integral:

$$\left\| \int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(s)) - f(s, y_\varepsilon(t_0 + iT)))\,\Delta s \right\| \le \int_{t_0+(i-1)T}^{t_0+iT} L\|y_\varepsilon(s) - y_\varepsilon(t_0 + iT)\|\,\Delta s \le \varepsilon LMT^2$$

$$\left\| \int_{t_0+(i-1)T}^{t_0+iT} (f^0(y_\varepsilon(s)) - f^0(y_\varepsilon(t_0 + iT)))\,\Delta s \right\| \le \int_{t_0+(i-1)T}^{t_0+iT} L\|y_\varepsilon(s) - y_\varepsilon(t_0 + iT)\|\,\Delta s \le \varepsilon LMT^2$$

Collecting these results together gives

$$\left\| \int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\,\Delta s \right\| \le 2\varepsilon LMT^2$$

for every $i \in \{1, \dots, m\}$. Using the fact that $mT \le t - t_0 \le d/\varepsilon$, we obtain

$$\left\| \sum_{i=1}^{m} \int_{t_0+(i-1)T}^{t_0+iT} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\, \Delta s \right\| \le 2m\varepsilon LMT^2 \le 2LMTd.$$

Combining this estimate with equality (2), we see that

$$\left\| \int_{t_0}^{t} (f(s, y_\varepsilon(s)) - f^0(y_\varepsilon(s)))\, \Delta s \right\| \le 2LMTd + 2MT.$$

Substituting this result into (1), we obtain

$$\|x_\varepsilon(t) - y_\varepsilon(t)\| \le \varepsilon(B + 2LMTd + 2MT) + \varepsilon L \int_{t_0}^{t} \|x_\varepsilon(s) - y_\varepsilon(s)\|\, \Delta s + \varepsilon^2 M(t - t_0).$$

It follows from Theorem 2.4 with $\alpha = \varepsilon(B + 2LMTd + 2MT)$, $\beta = \varepsilon^2 M$, $\gamma = \varepsilon L$ that

$$\|x_\varepsilon(t) - y_\varepsilon(t)\| \le \varepsilon \left(B + 2LMTd + 2MT + \frac{M}{L} \right) e_{L\varepsilon}(t, t_0) - \frac{\varepsilon M}{L}.$$

Finally, Theorem 2.5 gives

$$\begin{aligned} \|x_\varepsilon(t) - y_\varepsilon(t)\| &\le \varepsilon \left(B + 2LMTd + 2MT + \frac{M}{L} \right) e^{\varepsilon L(t-t_0)} - \frac{\varepsilon M}{L} \\ &\le \varepsilon \left(\left(B + 2LMTd + 2MT + \frac{M}{L} \right) e^{Ld} - \frac{M}{L} \right) = C\varepsilon, \end{aligned}$$

where C is a constant independent of ε. □

The set $U \subset \mathbb{R}^n$ from the assumptions can be a region of phase space that we are interested in. If f and g are bounded in U, we can choose the constant $d > 0$ so that the solutions x, y remain in U for every $\varepsilon \in (0, \varepsilon_0]$ and $t \in [t_0, t_0 + d/\varepsilon]_{\mathbb{T}}$. For example, let $x_0 \in \mathbb{R}^n$, $r > 0$ and $U = \{x \in \mathbb{R}^n;\ \|x - x_0\| \le r\}$. Assume that $\|f(t, x)\| \le M$ and $\|g(t, x, \varepsilon)\| \le M$ for every $x \in U$, $t \in [t_0, \infty)_{\mathbb{T}}$, $\varepsilon \in (0, \varepsilon_0]$. Consider the solution of

$$x^\Delta(t) = \varepsilon f(t, x(t)) + \varepsilon^2 g(t, x(t), \varepsilon), \quad x(t_0) = x_0$$

corresponding to $\varepsilon \in (0, \varepsilon_0]$. As long as x stays in U, its Δ-derivative is bounded by $\varepsilon M + \varepsilon^2 M \le \varepsilon M(1 + \varepsilon_0)$. Thus if we take $d < \frac{r}{M(1+\varepsilon_0)}$, then x cannot leave the ball U during the time interval $[t_0, t_0 + d/\varepsilon]_{\mathbb{T}}$. The same assertion is true for y, whose Δ-derivative is bounded by εM.

Alternatively, we can start by choosing the constant $d > 0$ and finding $\varepsilon_0 > 0$ and $U \subset \mathbb{R}^n$ such that $x, y \in U$ for every $\varepsilon \in (0, \varepsilon_0]$ and $t \in [t_0, t_0 + d/\varepsilon]_{\mathbb{T}}$.

The following example demonstrates the averaging method on a simple dynamic equation. We have deliberately chosen an equation which can be solved analytically without averaging, so that we can compare the solutions of the original and averaged equation.

Example 3.2. Consider the initial-value problem

$$x^\Delta(t) = \varepsilon(\sin(t\pi) + 1)x(t), \quad x(0) = x_0$$

on the time scale $\mathbb{T} = \bigcup_{k=0}^{\infty} [2k, 2k + 1]$. A similar equation was used in [4] to describe the growth of a plant population; in our case, the growth coefficient varies during the season instead of being constant. This equation has the form $x^\Delta(t) = \varepsilon f(t, x(t))$, where $f(t, x) = (\sin(t\pi) + 1)x$ is a 2-periodic function in the first argument. It is clear that f is continuous and bounded on $[t_0, \infty)_{\mathbb{T}} \times U$ whenever U is an arbitrary bounded set. Moreover, $\|f(t,x) - f(t,y)\| \leq 2\|x - y\|$ for every $t \in \mathbb{T}$ and $x,\ y \in \mathbb{R}^n$, which means that f is Lipschitz-continuous in the second argument. The exact solution of the dynamic equation is $x(t) = x_0 e_{\varepsilon(\sin(t\pi)+1)}(t, 0)$, i.e.

$$x(t) = x(2k) \exp\left(\int_{2k}^{t} \varepsilon(\sin(u\pi) + 1)\,\mathrm{d}u\right) = x(2k) \exp\left(\varepsilon\left(\frac{1 - \cos \pi t}{\pi} + t - 2k\right)\right),$$

$$x(2k + 2) = (1 + \varepsilon)x(2k + 1) = (1 + \varepsilon) \exp\left(\varepsilon(2/\pi + 1)\right)x(2k)$$

for every nonnegative integer k and every $t \in [2k, 2k + 1]$.

According to the averaging theorem, a good approximation of the exact solution can be obtained by considering the averaged equation $y^\Delta(t) = \varepsilon f^0(y(t))$, where

$$f^0(y) = \frac{1}{2}\int_0^2 f(t,y)\,\Delta t = \frac{1}{2}\left(\int_0^1 f(t,y)\,\mathrm{d}t + f(1,y)\right) = \frac{1}{2}\left(\int_0^1 (\sin(t\pi) + 1)y\,\mathrm{d}t + y\right) = (1 + 1/\pi)y.$$

Its solution is

$$y(t) = y(2k) \exp(\varepsilon(1 + 1/\pi)(t - 2k)),$$

$$y(2k + 2) = (1 + \varepsilon + \varepsilon/\pi)y(2k + 1) = (1 + \varepsilon + \varepsilon/\pi) \exp(\varepsilon(1 + 1/\pi)y(2k))$$

for every nonnegative integer k and every $t \in [2k, 2k + 1]$.

4 Existence of periodic solutions

We now present a generalization of the classical theorem which says that if the averaged equation

$$y'(t) = \varepsilon f^0(y(t))$$

has an equilibrium at p_0, then the original equation has a periodic solution near p_0 for all sufficiently small values of ε (see e.g. section 11.8 in [17]). The proof will be based on the so-called "near-identity transform": Given a pair of functions $x : [a, b]_{\mathbb{T}} \to \mathbb{R}^n$ and $u : [a, b]_{\mathbb{T}} \times \mathbb{R}^n \to \mathbb{R}^n$, the near-identity transform introduces a new function $z : [a, b]_{\mathbb{T}} \to \mathbb{R}^n$ given by the implicit formula

$$x(t) = z(t) + \varepsilon u(t, z(t)).$$

This is always possible for ε small enough, although many texts devoted to averaging and its applications do not discuss the validity of this transform. We provide a brief justification following the approach from [16, Lemma 2.8.3]. The symbol $B_r(p)$ will be used here and throughout the rest of the paper to denote the open ball $\{x \in \mathbb{R}^n;\ \|x - p\| < r\}$.

Definition 4.1. Assume that $X \subset Y \subset \mathbb{R}^n$. We say that X is an interior subset of Y, if there exists an $\varepsilon > 0$ such that if $x \in X$ and $\|y - x\| < \varepsilon$, then $y \in Y$.

Lemma 4.2. *Let $\mathbb{T}$ be a time scale, $U \subset \mathbb{R}^n$ an open set, $D \subset \mathbb{R}^n$ an interior bounded subset of U. Consider a continuous function $u : [a, b]_{\mathbb{T}} \times U \to \mathbb{R}^n$, which is continuously Δ-differentiable in the first argument and whose partial derivatives $\frac{\partial u}{\partial x_1}, \dots, \frac{\partial u}{\partial x_n}$ exist and are continuous in $[a, b]_{\mathbb{T}} \times U$. Then there exists a number $\varepsilon_0 > 0$ and a unique function $\varphi : [a, b]_{\mathbb{T}} \times (-\varepsilon_0, \varepsilon_0) \times D \to U$ which satisfies*

$$x = \varphi(t, \varepsilon, x) + \varepsilon u(t, \varphi(t, \varepsilon, x))$$

on its domain, is continuously differentiable with respect to ε and x, and Δ-differentiable with respect to t.

Proof. Since D is an interior subset of U, there exists a $\delta > 0$ such that if $x \in D$ and $\|y - x\| < \delta$, then $y \in U$. Let

$$F(t, \varepsilon, x, z) = z + \varepsilon u(t, z) - x \tag{3}$$

for every $t \in [a, b]_{\mathbb{T}}$, $\varepsilon \in \mathbb{R}$, $x \in \mathbb{R}^n$, $z \in U$. This function is continuously differentiable with respect to x, z and ε, and continuously Δ-differentiable with respect to t. For every $\tau \in [a, b]_{\mathbb{T}}$ and every $x_0 \in \overline{D}$, we see that $F(\tau, 0, x_0, x_0) = 0$ and $\frac{\partial F}{\partial z}(\tau, 0, x_0, x_0) = I$. According to Theorem 2.10 with $l = n + 2$, $\mathbb{T}_1 = [a, b]_{\mathbb{T}}$ and $\mathbb{T}_2 = \cdots = \mathbb{T}_{n+2} = \mathbb{R}$, there exist positive numbers $\delta_1(\tau, x_0)$, $\delta_2(\tau, x_0)$, $\delta_3(\tau, x_0)$, and a unique continuous function $\varphi : (B_{\delta_1(\tau,x_0)}(\tau) \cap [a, b]_{\mathbb{T}}) \times B_{\delta_2(\tau,x_0)}(0) \times B_{\delta_3(\tau,x_0)}(x_0) \to \mathbb{R}^n$ which satisfies

$$F(t, \varepsilon, x, \varphi(t, \varepsilon, x)) = 0 \tag{4}$$

in its domain; note that by (3), this is equivalent to

$$x = \varphi(t, \varepsilon, x) + \varepsilon u(t, \varphi(t, \varepsilon, x)). \tag{5}$$

We know that $\varphi(\tau, 0, x_0) = x_0$; by continuity of φ, we can assume that $\delta_1(\tau, x_0)$, $\delta_2(\tau, x_0)$, $\delta_3(\tau, x_0)$ are small enough to ensure that $\|\varphi(t, \varepsilon, x) - x\| < \delta$ for every t, ε, x, i.e. $\varphi(t, \varepsilon, x) \in U$ whenever $x \in D$.

Since $[a, b]_{\mathbb{T}} \times \overline{D}$ is a compact set, it can be covered by a finite number of neighborhoods of the form $(B_{\delta_1(\tau,x_0)}(\tau) \cap [a, b]_{\mathbb{T}}) \times B_{\delta_3(\tau,x_0)}(x_0)$, where $\tau \in [a, b]_{\mathbb{T}}$ and $x_0 \in \overline{D}$. Let ε_0 be the minimum of the corresponding values of $\delta_2(\tau, x_0)$. Then for every $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$, $t \in [a, b]_{\mathbb{T}}$ and $x \in D$, there is a unique $\varphi(t, \varepsilon, x) \in U$ such that (5) is satisfied.

Finally, Theorem 2.10 guarantees that φ has partial derivatives with respect to $x_1, \ldots, x_n$, ε and partial Δ-derivative with respect to t; differentiation of (4) shows that the partial derivatives with respect to $x_1, \ldots, x_n$, ε are continuous. □

We also need the following statement concerning existence of solutions and differentiability with respect to parameters, which is a consequence of Theorem 3.1 in [7].

Theorem 4.3. *Let $p_0 \in \mathbb{R}^n$, $r > 0$, $\varepsilon_0 > 0$. Assume that $f : [a, b]_{\mathbb{T}} \times B_r(p_0) \to \mathbb{R}^n$ and $g : [a, b]_{\mathbb{T}} \times B_r(p_0) \times (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}^n$ are continuous, and that $\frac{\partial f}{\partial x}$, $\frac{\partial g}{\partial x}$, $\frac{\partial g}{\partial \varepsilon}$ exist and are continuous in their domains. Then there exist $\varepsilon_1 \in (0, \varepsilon_0)$ and $\delta \in (0, r)$ such that for every $\varepsilon \in (-\varepsilon_1, \varepsilon_1)$ and every $x_0 \in B_\delta(p_0)$, the initial-value problem*

$$x^\Delta(t) = \varepsilon f(t, x(t)) + \varepsilon^2 g(t, x(t), \varepsilon), \quad x(t_0) = x_0$$

has a unique solution $t \mapsto x(t, x_0, \varepsilon)$ defined on $[a, b]_{\mathbb{T}}$. For every $t \in [a, b]_{\mathbb{T}}$, $x(t, x_0, \varepsilon)$ is continuously differentiable with respect to x_0 and ε.

We now proceed to the promised theorem on the existence of periodic solutions.

Theorem 4.4. *Consider a T-periodic time scale $\mathbb{T}$, $t_0 \in \mathbb{T}$, $p_0 \in \mathbb{R}^n$, $r > 0$, $\varepsilon_0 > 0$, and a pair of continuous functions $f : [t_0, \infty)_{\mathbb{T}} \times B_r(p_0) \to \mathbb{R}^n$, $g : [t_0, \infty)_{\mathbb{T}} \times B_r(p_0) \times (-\varepsilon_0, \varepsilon_0) \to \mathbb{R}^n$. Let*

$$f^0(x) = \frac{1}{T} \int_{t_0}^{t_0+T} f(t, x)\, \Delta t, \quad x \in B_r(p_0).$$

Assume that the following conditions are satisfied:

1. *$\frac{\partial f}{\partial x_i}$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}$ exist and are continuous in $[t_0, \infty)_{\mathbb{T}} \times B_r(p_0)$ for every $i, j \in \{1, \ldots, n\}$.*
2. *$\frac{\partial g}{\partial \varepsilon}$ and $\frac{\partial g}{\partial x}$ exist and are continuous in $[t_0, \infty)_{\mathbb{T}} \times B_r(p_0) \times (-\varepsilon_0, \varepsilon_0)$.*
3. *f and g are T-periodic in the first argument.*

4. $f^0(p_0)=0$ *and* $\det\frac{\partial f^0}{\partial x}(p_0)\neq 0$.

Then there exist numbers $\varepsilon_1\in(0,\varepsilon_0)$, $C>0$ *and a continuous function* $p:[-\varepsilon_1,\varepsilon_1]\to B_r(p_0)$ *such that* $p(0)=p_0$ *and for every* $\varepsilon\in[-\varepsilon_1,\varepsilon_1]$, *the initial-value problem*

$$x^\Delta(t)=\varepsilon f(t,x(t))+\varepsilon^2 g(t,x(t),\varepsilon),\ \ x(t_0)=p(\varepsilon)$$

has a unique solution $t\mapsto x(t,p(\varepsilon),\varepsilon)$ *defined on* $[t_0,\infty)_{\mathbb{T}}$, *which is* T*-periodic and satisfies*

$$\|x(t,p(\varepsilon),\varepsilon)-p_0\|\leq C|\varepsilon|,\ \ t\in[t_0,\infty)_{\mathbb{T}}.$$

Proof. Let

$$u(t,x)=\int_{t_0}^t (f(s,x)-f^0(x))\,\Delta s,\quad x\in B_r(p_0),\ t\in[t_0,\infty)_{\mathbb{T}}.$$

It follows from the definition of f^0 and Lemma 2.3 that the integral in the definition of u is zero when taken over an arbitrary interval of length T, and thus u is T-periodic in the first argument. Note that u is always Δ-differentiable with respect to t. According to Theorem 2.6, $\frac{\partial u}{\partial x}$ exists and is continuous in $[t_0,\infty)_{\mathbb{T}}\times B_r(p_0)$.

Choose an arbitrary $\delta_2\in(0,r)$. According to Theorem 4.3, there exist numbers $\varepsilon_2\in(0,\varepsilon_0)$ and $\delta_3\in(0,r)$ such that for every $\varepsilon\in(-\varepsilon_2,\varepsilon_2)$ and every $x_0\in B_{\delta_3}(p_0)$, the initial-value problem

$$x^\Delta(t)=\varepsilon f(t,x(t))+\varepsilon^2 g(t,x(t),\varepsilon),\ \ x(t_0)=x_0$$

has a solution $t\mapsto x(t,x_0,\varepsilon)$, which is defined on $[t_0,t_0+T]_{\mathbb{T}}$, takes values in $B_{\delta_2}(p_0)$, and $x(t,x_0,\varepsilon)$ is continuously differentiable with respect to ε and x_0. Let $\delta_1\in(\delta_2,r)$. By Lemma 4.2, there exists an $\varepsilon_3\in(0,\varepsilon_2)$ such that if $x\in B_{\delta_2}(p_0)$, $t\in[t_0,t_0+T]_{\mathbb{T}}$, and $\varepsilon\in(-\varepsilon_3,\varepsilon_3)$, there is a $z=\varphi(t,\varepsilon,x)\in B_{\delta_1}(p_0)$ such that

$$x=z+\varepsilon u(t,z).$$

We also know from Lemma 4.2 that the near-identity transform $z=\varphi(t,\varepsilon,x)$ is continuously differentiable with respect to x and Δ-differentiable with respect to t. Given a fixed $\varepsilon\in(-\varepsilon_3,\varepsilon_3)$ and a Δ-differentiable function $x:[t_0,t_0+T]_{\mathbb{T}}\to B_{\delta_2}$, we can let $z(t)=\varphi(t,\varepsilon,x(t))$ and thus obtain a function $z:[t_0,t_0+T]_{\mathbb{T}}\to B_{\delta_1}(p_0)$ such that

$$x(t)=z(t)+\varepsilon u(t,z(t)).$$

Note that z is Δ-differentiable; this is clear at right-scattered points, and at a right-dense point t, we have

$$z^\Delta(t)=\varphi^\Delta(t,\varepsilon,x(t))+\frac{\partial\varphi}{\partial x}(t,\varepsilon,x(t))x^\Delta(t).$$

Now, assume that the function x satisfies $x^\Delta(t)=\varepsilon f(t,x(t))+\varepsilon^2 g(t,x(t),\varepsilon)$. If t is right-dense, we obtain

$$\varepsilon f(t,x(t))+\varepsilon^2 g(t,x(t),\varepsilon)=x^\Delta(t)=z^\Delta(t)+\varepsilon u^\Delta(t,z(t))+\varepsilon\frac{\partial u}{\partial x}(t,z(t))z^\Delta(t).$$

Using the obvious identity $u^\Delta(t,z(t))=f(t,z(t))-f^0(z(t))$, rearranging the terms and substituting $x(t)=z(t)+\varepsilon u(t,z(t))$, we conclude that

$$\left(I+\varepsilon\frac{\partial u}{\partial x}(t,z(t))\right)z^\Delta(t)=\varepsilon f^0(z(t))-\varepsilon f(t,z(t))+\varepsilon f(t,z(t)+\varepsilon u(t,z(t)))+$$
$$+\varepsilon^2 g(t,z(t)+\varepsilon u(t,z(t)),\varepsilon).\qquad(6)$$

On the other hand, if t is right-scattered, we have

$$\varepsilon f(t,x(t))+\varepsilon^2 g(t,x(t),\varepsilon)=x^{\Delta}(t)=z^{\Delta}(t)+\varepsilon\frac{u(\sigma(t),z(\sigma(t)))-u(\sigma(t),z(t))}{\mu(t)}+$$

$$+\varepsilon\frac{u(\sigma(t),z(t))-u(t,z(t))}{\mu(t)}=z^{\Delta}(t)+\varepsilon\frac{u(\sigma(t),z(\sigma(t)))-u(\sigma(t),z(t))}{\mu(t)}+\varepsilon u^{\Delta}(t,z(t)). \tag{7}$$

The mean-value theorem for vector-valued functions (see e.g. [11], page 278) gives

$$u(\sigma(t),z(\sigma(t)))-u(\sigma(t),z(t))=\left(\int_0^1\frac{\partial u}{\partial x}(\sigma(t),(1-s)z(\sigma(t))+sz(t))\,\mathrm{d}s\right)(z(\sigma(t))-z(t)).$$

If we introduce the function $M:[t_0,t_0+T]_{\mathbb{T}}\times B_r(p_0)\times B_r(p_0)\to\mathbb{R}^n$ by the formula

$$M(t,a,b)=\int_0^1\frac{\partial u}{\partial x}(\sigma(t),(1-s)b+sa)\,\mathrm{d}s,$$

it follows that

$$\frac{u(\sigma(t),z(\sigma(t))-u(\sigma(t),z(t))}{\mu(t)}=M(t,z(t),z(\sigma(t)))z^{\Delta}(t).$$

Substituting this into (7), using the identity $u^{\Delta}(t,z(t))=f(t,z(t))-f^0(z(t))$, and rearranging the terms as in the right-dense case, we obtain

$$(I+\varepsilon M(t,z(t),z(\sigma(t))))z^{\Delta}(t)=\varepsilon f^0(z(t))-\varepsilon f(t,z(t))+\varepsilon f(t,z(t)+\varepsilon u(t,z(t)))+$$

$$+\varepsilon^2 g(t,z(t)+\varepsilon u(t,z(t)),\varepsilon). \tag{8}$$

Note that if t is right-dense, then $M(t,z(t),z(\sigma(t)))=\frac{\partial u}{\partial x}(t,z(t))$. Comparing this with equation (6), we see that equation (8) is in fact true for both right-dense and right-continuous points t.

M is bounded on $[t_0,t_0+T]_{\mathbb{T}}\times B_{\delta_1}(p_0)\times B_{\delta_1}(p_0)$, thus if $|\varepsilon|$ is sufficiently small, we have

$$(I+\varepsilon M(t,z(t),z(\sigma(t))))^{-1}=\sum_{k=0}^{\infty}(-1)^k\varepsilon^k M(t,z(t),z(\sigma(t)))^k.$$

Consequently, it follows from (8) that

$$z^{\Delta}(t)=\left(\sum_{k=0}^{\infty}(-1)^k\varepsilon^k M(t,z(t),z(\sigma(t)))^k\right)\big(\varepsilon f^0(z(t))-\varepsilon f(t,z(t))+\varepsilon f(t,z(t)+\varepsilon u(t,z(t)))+$$

$$+\varepsilon^2 g(t,z(t)+\varepsilon u(t,z(t)),\varepsilon)\big). \tag{9}$$

Using the mean-value theorem again, we obtain

$$f(t,z+\varepsilon u(t,z))-f(t,z)=\left(\int_0^1\frac{\partial f}{\partial x}(t,z+\varepsilon su(t,z))\,\mathrm{d}s\right)\varepsilon u(t,z).$$

Thus, if we let

$$h(t,z,\varepsilon)=\left(\int_0^1\frac{\partial f}{\partial x}(t,z+\varepsilon su(t,z))\,\mathrm{d}s\right)u(t,z),$$

then (9) can be rewritten as

$$z^{\Delta}(t)=\left(\sum_{k=0}^{\infty}(-1)^k\varepsilon^k M(t,z(t),z(\sigma(t)))^k\right)\left(\varepsilon f^0(z(t))+\varepsilon^2\big(h(t,z(t),\varepsilon)+g(t,z(t)+\varepsilon u(t,z(t)),\varepsilon)\big)\right).$$

After collecting all terms of order ε^2 together, we obtain

$$z^\Delta(t) = \varepsilon f^0(z(t)) + \varepsilon^2 R(t, z(t), z(\sigma(t)), \varepsilon),$$

where

$$R(t, z, w, \varepsilon) = f^0(z)\left(\sum_{k=1}^{\infty}(-1)^k\varepsilon^{k-1}M(t,z,w)^k\right) +$$

$$+\big(h(t,z,\varepsilon) + g(t, z+\varepsilon u(t,z),\varepsilon)\big)\left(\sum_{k=0}^{\infty}(-1)^k\varepsilon^k M(t,z,w)^k\right).$$

Note that $z(t_0) = x(t_0)$. Thus we have proved that if $|\varepsilon|$ is sufficiently small and $z_0 \in B_{\delta_3}(p_0)$, the initial-value problem

$$z^\Delta(t) = \varepsilon f^0(z(t)) + \varepsilon^2 R(t, z(t), z(\sigma(t)), \varepsilon), \quad z(t_0) = z_0 \tag{10}$$

has a solution $t \mapsto z(t, z_0, \varepsilon)$ defined on $[t_0, t_0+T]_{\mathbb{T}}$, which takes values in $B_{\delta_1}(p_0)$ and is given by the implicit formula

$$x(t, z_0, \varepsilon) = z(t, z_0, \varepsilon) + \varepsilon u(t, z(t, z_0, \varepsilon)).$$

Since the near-identity transform $x(t, z_0, \varepsilon) \mapsto z(t, z_0, \varepsilon)$ is continuously differentiable with respect to ε and z_0, it follows that $z(t, z_0, \varepsilon)$ is also continuously differentiable with respect to ε and z_0. Let

$$F(z_0, \varepsilon) = \int_{t_0}^{t_0+T} f^0(z(t, z_0, \varepsilon)) + \varepsilon R(t, z(t, z_0, \varepsilon), z(\sigma(t), z_0, \varepsilon), \varepsilon)\Delta t.$$

Note that by (10), $z(t_0, z_0, \varepsilon) = z(t_0+T, z_0, \varepsilon)$ if and only if $F(z_0, \varepsilon) = 0$.

If $\varepsilon = 0$, then $z(t, z_0, \varepsilon) = z_0$ for every $t \in [t_0, t_0+T]_{\mathbb{T}}$, $F(z_0, 0) = Tf^0(z_0)$, and consequently

$$F(p_0, 0) = Tf^0(p_0) = 0,$$

$$\det\frac{\partial F}{\partial z_0}(p_0, 0) = T\det\frac{\partial f^0}{\partial x}(p_0) \neq 0.$$

An inspection of the definition of R reveals that this function is continuously differentiable in the second, third and fourth argument. Since both $z(t, z_0, \varepsilon)$ and $z(\sigma(t), z_0, \varepsilon)$ are continuously differentiable with respect to ε and z_0, we conclude that F is also continuously differentiable with respect to ε and z_0.

According to the implicit function theorem, there exists a number $\varepsilon_4 \in (0, \varepsilon_0)$ and a continuous mapping $p : (-\varepsilon_4, \varepsilon_4) \to B_r(p_0)$ such that $F(p(\varepsilon), \varepsilon) = 0$ for every $\varepsilon \in (-\varepsilon_4, \varepsilon_4)$, i.e. $z(t_0, p(\varepsilon), \varepsilon) = z(t_0+T, p(\varepsilon), \varepsilon)$. Consequently, $x(t_0, p(\varepsilon), \varepsilon) = x(t_0+T, p(\varepsilon), \varepsilon)$. Thus if x is extended T-periodically to $[t_0, \infty)_{\mathbb{T}}$, we obtain a T-periodic solution of $x^\Delta(t) = \varepsilon f(t, x(t)) + \varepsilon^2 g(t, x(t), \varepsilon)$.

The implicit function theorem guarantees that p is differentiable at 0, and thus there exist numbers $B > 0$, $\varepsilon_1 \in (0, \varepsilon_2)$ such that $\|p(\varepsilon) - p_0\| \le B|\varepsilon|$ if $|\varepsilon| \le \varepsilon_1$. We let $d = \varepsilon_1 T$. By the periodic averaging theorem (the Lipschitz-continuity assumption follows from the continuity of $\frac{\partial f}{\partial x}$), there exists a number $C > 0$ such that

$$\|x(t, p(\varepsilon), \varepsilon) - p_0\| \le C\varepsilon$$

for every $t \in [t_0, t_0+T]_{\mathbb{T}}$ and every $\varepsilon \in (0, \varepsilon_1]$. The solution x is T-periodic, and therefore the last estimate is in fact valid for every $t \in [t_0, \infty)_{\mathbb{T}}$. We still need to obtain the corresponding estimate for $\varepsilon < 0$. This can be done by considering the functions $\tilde{f}(t, x) = -f(t, x)$, $\tilde{g}(t, x, \varepsilon) = g(t, x, -\varepsilon)$, $\tilde{p}(\varepsilon) = p(-\varepsilon)$, $\tilde{x}(t, x_0, \varepsilon) = x(t, x_0, -\varepsilon)$ and applying the periodic averaging theorem to the initial-value problem

$$\tilde{x}^\Delta(t) = \varepsilon\tilde{f}(t, \tilde{x}(t))) + \varepsilon^2\tilde{g}(t, \tilde{x}(t), \varepsilon), \quad \tilde{x}(t_0) = \tilde{p}(\varepsilon)$$

with $\varepsilon > 0$. □

Example 4.5. Consider the time scale $\mathbb{T} = \mathbb{Z}$ and the dynamic equation

$$x^{\Delta}(t) = \varepsilon(1 - x(t) + (-1)^t), \quad t \in \{0, 1, 2, \dots\},$$

whose right-hand side is 2-periodic in t. The corresponding averaged equation is $y^{\Delta}(t) = \varepsilon f^0(y(t))$, where $f^0(x) = 1 - x$. It has an equilibrium solution $y(t) = p_0 = 1$. Moreover, $\frac{\partial f^0}{\partial x}(p_0) = -1$. Thus the previous theorem guarantees that the original dynamic equation has a 2-periodic solution near p_0 whenever $|\varepsilon|$ is sufficiently small. Indeed, the equation is so simple that we can solve it analytically and look for a 2-periodic solution; the result is $x(t) = 1 + (-1)^t \varepsilon/(\varepsilon - 2)$. Using the notation of Theorem 4.4, we have $p(\varepsilon) = x(0) = 1 + \varepsilon/(\varepsilon - 2)$. If we restrict ourselves to $\varepsilon \in [-1, 1]$, we have $|\varepsilon - 2| \geq 1$ and $|x(t) - 1| = |\varepsilon/(\varepsilon - 2)| \leq |\varepsilon|$ for every $t \in \{0, 1, 2, \dots\}$.

5 Nonperiodic averaging

Consider an interval $I \subset \mathbb{R}$, a set $B \subset \mathbb{R}^n$, and a function $F : B \times I \to \mathbb{R}^n$. A function $x : I \to B$ is called a solution of the generalized ordinary differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = DF(x, t),$$

whenever

$$x(s_2) - x(s_1) = \int_{s_1}^{s_2} DF(x(\tau), t)$$

for each pair of values $s_1, s_2 \in I$. The integral on the right-hand side is the Kurzweil integral defined as the limit of the integral sums

$$\sum_{j=1}^{k} \big(F(\tau_j, \alpha_j) - F(\tau_j, \alpha_{j-1})\big)$$

with respect to δ-fine partitions $a = \alpha_0 \leq \alpha_1 \leq \dots \leq \alpha_k = b$, $\tau_i \in [\alpha_{i-1}, \alpha_i]$, $i = 1, \dots, k$ (see [12] for more details). An important special case of this integral is the Kurzweil-Stieltjes integral $\int_a^b f(s)\,\mathrm{d}g(s)$, which is obtained from a pair of functions $f : [a, b] \to \mathbb{R}^n$ and $g : [a, b] \to \mathbb{R}$ by setting $F(\tau, t) = f(\tau)g(t)$. Generalized ordinary differential equations were introduced by Jaroslav Kurzweil in 1957 (see the paper [9]). It is now known that this concept includes not only ordinary differential equations, but also differential equations with impulses, measure differential equations, retarded functional differential equations, dynamic equations on time scales etc.

In this section, we use the existing averaging theorem for generalized ordinary differential equations to derive a nonperiodic averaging theorem for dynamic equations on time scales.

In the following text, we use the same notation as in [13]: if $t \leq \sup \mathbb{T}$, we let

$$t^* = \inf\{s \in \mathbb{T}; s \geq t\}.$$

(Note that t^* is different from $\sigma(t) = \inf\{s \in \mathbb{T}; s > t\}$.) Since $\mathbb{T}$ is a closed set, we have $t^* \in \mathbb{T}$. Further, let

$$\mathbb{T}^* = \begin{cases} (-\infty, \sup \mathbb{T}] & \text{if } \sup \mathbb{T} < \infty, \\ (-\infty, \infty) & \text{otherwise.} \end{cases}$$

Given a function $f : \mathbb{T} \to \mathbb{R}^n$, we define a function $f^* : \mathbb{T}^* \to \mathbb{R}^n$ by

$$f^*(t) = f(t^*), \quad t \in \mathbb{T}^*.$$

Similarly, given a set $B \subset \mathbb{R}^n$ and a function $f : B \times \mathbb{T} \to \mathbb{R}^n$, we let

$$f^*(x, t) = f(x, t^*), \quad x \in B, \ t \in \mathbb{T}^*.$$

Finally, a function $f : B \times \mathbb{T} \to \mathbb{R}^n$ is called rd-continuous, if the function $t \mapsto f(x(t), t)$ is rd-continuous whenever $x : \mathbb{T} \to B$ is a continuous function.

The following theorem describes a one-to-one correspondence between the solutions of a dynamic equation and the solutions of a certain generalized ordinary differential equation. The statement is a special case of Theorem 12 from [13].

Theorem 5.1. *Let B be a bounded subset of $\mathbb{R}^n$ and $f : B \times \mathbb{T} \to \mathbb{R}^n$ a bounded rd-continuous function, which is Lipschitz-continuous in the first variable. If $x : \mathbb{T} \to B$ is a solution of*

$$x^{\Delta}(t) = f(x(t), t), \quad t \in \mathbb{T}, \tag{11}$$

then $x^* : \mathbb{T}^* \to B$ is a solution of the generalized ordinary differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = DF(x, t), \quad t \in \mathbb{T}^*, \tag{12}$$

where

$$F(x, t) = \int_{t_0}^{t} f(x, s^*) \, \mathrm{d}g(s), \quad x \in B, \quad t \in \mathbb{T}^*,$$

$t_0 \in \mathbb{T}$ is an arbitrary fixed number, and $g(s) = s^*$ for every $s \in \mathbb{T}^*$. Moreover, every solution $y : \mathbb{T}^* \to B$ of (12) has the form $y = x^*$, where $x : \mathbb{T} \to B$ is a solution of (11).

We proceed to the averaging theorem for generalized ordinary differential equations (the statement as well as its proof can be found in [12], Theorem 8.12).

Theorem 5.2. *Consider a number $r > 0$ and a function $F : B_r(0) \times [0, \infty) \to \mathbb{R}^n$ such that the following conditions are satisfied:*

1. *There exists a nondecreasing function $h : [0, \infty) \to \mathbb{R}$ such that*

   $$\|F(x, t_2) - F(x, t_1)\| \le |h(t_2) - h(t_1)|$$

   for every $x \in B_r(0)$ and $t_1, t_2 \in [0, \infty)$.

2. *There exists a continuous increasing function $\omega : [0, \infty) \to \mathbb{R}$ such that $\omega(0) = 0$ and*

   $$\|F(x, t_2) - F(x, t_1) - F(y, t_2) + F(y, t_1)\| \le \omega(\|x - y\|) |h(t_2) - h(t_1)|$$

   for every $x, y \in B_r(0)$ and $t_1, t_2 \in [0, \infty)$.

3. *There exists a number $C \in \mathbb{R}$ such that for every $a \in [0, \infty)$*

   $$\limsup_{r \to \infty} \frac{h(a + r) - h(a)}{r} \le C.$$

4. *There exists a function $F_0 : B_r(0) \to \mathbb{R}^n$ such that*

   $$\lim_{r \to \infty} \frac{F(x, r)}{r} = F_0(x), \quad x \in B_r(0).$$

Let $x_0 \in B_r(0)$. Assume that the equation

$$y'(t) = F_0(y(t)), \quad y(0) = x_0$$

has a unique solution $y : [0,\infty) \to \mathbb{R}^n$*, which is contained in an interior subset of* $B_r(0)$*. Then for every* $\mu > 0$ *and* $d > 0$ *there is an* $\varepsilon_0 > 0$ *such that for every* $\varepsilon \in (0, \varepsilon_0)$*, the generalized ordinary differential equation*

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = D(\varepsilon F(x,t)), \;\; x(0) = x_0$$

has a solution $x_\varepsilon : [0, d/\varepsilon] \to \mathbb{R}^n$*, the ordinary differential equation*

$$y'(t) = \varepsilon F_0(y(t)), \;\; y(0) = x_0$$

has a solution $y_\varepsilon : [0, d/\varepsilon] \to \mathbb{R}^n$*, and*

$$\|x_\varepsilon(t) - y_\varepsilon(t)\| < \mu \text{ for every } t \in [0, d/\varepsilon].$$

With the help of the previous two statements, we prove the following averaging theorem for dynamic equations on time scales.

Theorem 5.3. *Let* $\mathbb{T}$ *be a time scale with* $\sup \mathbb{T} = \infty$ *and* $\limsup_{t\to\infty} \mu(t)/t < \infty$*,* $t_0 \in \mathbb{T}$*,* $p \in \mathbb{R}^n$*,* $r > 0$*. Assume that* $f : B_r(p) \times [t_0,\infty)_{\mathbb{T}} \to \mathbb{R}^n$ *is bounded, rd-continuous, and Lipschitz-continuous in the first variable. Moreover, assume there exists a function* $f^0 : B_r(p) \to \mathbb{R}^n$ *such that*

$$\lim_{T\to\infty} \frac{1}{T} \int_{t_0}^{t_0+T} f^*(x,s)\,\mathrm{d}g(s) = f^0(x), \;\; x \in B_r(p),$$

where $g(s) = s^*$ *for every* $s \in [t_0,\infty)$*. Let* $x_0 \in B_r(p)$*. If the equation*

$$y'(t) = f^0(y(t)), \;\; y(t_0) = x_0$$

has a unique solution $y : [t_0,\infty) \to \mathbb{R}^n$*, which is contained in an interior subset of* $B_r(p)$*, then for every* $\mu > 0$ *and* $d > 0$ *there is an* $\varepsilon_0 > 0$ *such that for every* $\varepsilon \in (0,\varepsilon_0)$*, the dynamic equation*

$$x^\Delta(t) = \varepsilon f(x(t),t), \;\; x(t_0) = x_0$$

has a solution $x_\varepsilon : [t_0, t_0 + d/\varepsilon]_{\mathbb{T}} \to \mathbb{R}^n$*, the ordinary differential equation*

$$y'(t) = \varepsilon f^0(y(t)), \;\; y(t_0) = x_0$$

has a solution $y_\varepsilon : [t_0, t_0 + d/\varepsilon] \to \mathbb{R}^n$*, and*

$$\|x_\varepsilon(t) - y_\varepsilon(t)\| < \mu \text{ for every } t \in [t_0, t_0 + d/\varepsilon]_{\mathbb{T}}.$$

Proof. Assume without loss of generality that $t_0 = 0$; otherwise, consider a shifted problem with time scale $\tilde{\mathbb{T}} = \{t - t_0;\ t \in \mathbb{T}\}$ and right-hand side $\tilde{f}(x,t) = f(x, t_0 + t)$. Similarly, we can assume that $p = 0$. By the assumptions, there exist numbers $m, l \in \mathbb{R}^+$ such that $\|f(x,t)\| \le m$ and $\|f(x,t) - f(y,t)\| \le l\|x-y\|$ for every $x, y \in B_r(p)$ and $t \in [t_0,\infty)_{\mathbb{T}}$. Let $h(t) = m \cdot g(t)$ for $t \in [0,\infty)$, $\omega(r) = \frac{l}{m} \cdot r$ for every $r \in [0,\infty)$, and

$$F(x,t) = \int_0^t f^*(x,s)\,\mathrm{d}g(s), \;\; x \in B_r(0), \;\; t \in [0,\infty).$$

When $0 \le t_1 \le t_2$ and $x \in B_r(0)$, we have

$$\|F(x,t_2) - F(x,t_1)\| = \left\| \int_{t_1}^{t_2} f^*(x,s)\,\mathrm{d}g(s) \right\| \le m(g(t_2) - g(t_1)) = h(t_2) - h(t_1).$$

Similarly, if $0 \le t_1 \le t_2$ and $x, y \in B_r(0)$, then

$$\|F(x,t_2) - F(x,t_1) - F(y,t_2) + F(y,t_1)\|$$

$$= \left\| \int_{t_1}^{t_2} f^*(x,s)\,\mathrm{d}g(s) - \int_{t_1}^{t_2} f^*(y,s)\,\mathrm{d}g(s) \right\| = \left\| \int_{t_1}^{t_2} (f^*(x,s) - f^*(y,s))\,\mathrm{d}g(s) \right\|$$
$$\le l\|x-y\|(g(t_2) - g(t_1)) = \omega(\|x-y\|)m(t_2^* - t_1^*) = \omega(\|x-y\|)(h(t_2) - h(t_1)).$$

Since $\limsup_{t\to\infty} \mu(t)/t < \infty$, there exists numbers $D > 0$ and $\tau \in \mathbb{T}$ such that $\mu(t)/t \le D$ for every $t \in [\tau, \infty)_{\mathbb{T}}$. It follows that if $t \in \mathbb{R}$ is such that $\rho(t^*) \ge \tau$, then

$$t^* \le t + \mu(\rho(t^*)) \le t + D\rho(t^*) \le t + Dt = t(D+1).$$

Thus for sufficiently large r we obtain

$$\frac{h(a+r) - h(a)}{r} = \frac{m((a+r)^* - a^*)}{r} \le \frac{m((a+r)(D+1) - a^*)}{r},$$

and consequently

$$\limsup_{r\to\infty} \frac{h(a+r) - h(a)}{r} \le \lim_{r\to\infty} \frac{m((a+r)(D+1) - a^*)}{r} = m(D+1).$$

It is also obvious that

$$\lim_{r\to\infty} \frac{F(x,r)}{r} = f^0(x), \quad x \in B_r(0),$$

and thus we see that F satisfies all four assumptions of Theorem 5.2. According to this theorem, given a $\mu > 0$ and $d > 0$, there is an $\varepsilon_0 > 0$ such that for every $\varepsilon \in (0, \varepsilon_0)$, the generalized ordinary differential equation

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = D(\varepsilon F(x,t)), \quad x(0) = 0$$

has a solution $x_\varepsilon : [0, d/\varepsilon] \to \mathbb{R}^n$, the ordinary differential equation

$$y'(t) = \varepsilon f^0(y(t)), \quad y(t_0) = x_0$$

has a solution $y_\varepsilon : [0, d/\varepsilon] \to \mathbb{R}^n$, and

$$\|x_\varepsilon(t) - y_\varepsilon(t)\| < \mu \text{ for every } t \in [0, d/\varepsilon].$$

To conclude the proof, it is sufficient to observe that, according to Theorem 5.1, the restriction of x_ε to $[0, d/\varepsilon]_{\mathbb{T}}$ coincides with the solution of the dynamic equation

$$x^\Delta(t) = \varepsilon f(x(t), t), \quad x(0) = x_0. \qquad \square$$

Our averaging theorem says that solutions of the original dynamic equation might be approximated by solutions of a certain autonomous differential equation, whose right-hand side is calculated with the help of the Kurzweil-Stieltjes integral. This integral might be difficult to evaluate, but as we now show, it is often possible to obtain the averaged function with the help of an ordinary Δ-integral. We need the following statement, which is a consequence of Theorem 5 in [13].

Theorem 5.4. *Consider a time scale $\mathbb{T}$ with $\sup \mathbb{T} = \infty$. Let $f : [t_0, \infty)_{\mathbb{T}} \to \mathbb{R}^n$ be a rd-continuous function. Choose an arbitrary $a \in [t_0, \infty)_{\mathbb{T}}$ and let*

$$\begin{aligned} F_1(t) &= \int_a^t f(s)\,\Delta s, \quad t \in [t_0, \infty)_{\mathbb{T}}, \\ F_2(t) &= \int_a^t f^*(s)\,\mathrm{d}g(s), \quad t \in [t_0, \infty), \end{aligned}$$

where $g(s) = s^*$ for every $s \in [t_0, \infty)$. Then $F_2(t) = F_1^*(t)$ for every $t \in [t_0, \infty)$; in particular, $F_2(t) = F_1(t)$ for every $t \in [t_0, \infty)_{\mathbb{T}}$.

Lemma 5.5. *Consider a time scale $\mathbb{T}$ such that $\sup \mathbb{T} = \infty$ and $\lim_{t\to\infty} \mu(t)/t = 0$, a number $t_0 \in \mathbb{T}$ and a function $H : [t_0, \infty)_{\mathbb{T}} \to \mathbb{R}^n$ such that*

$$\lim_{t\to\infty,\, t\in\mathbb{T}} \frac{H(t)}{t} = L.$$

Then

$$\lim_{t\to\infty} \frac{H^*(t)}{t} = L.$$

Proof. Given an arbitrary $\varepsilon > 0$, we can find a $T > \max(t_0, 0)$ such that for every $t \geq T$ we have

$$\left\| \frac{H(t)}{t} - L \right\| < \varepsilon,$$

$$\frac{\mu(t)}{t} < \varepsilon.$$

Then

$$\left\| \frac{H^*(t)}{t^*} - L \right\| < \varepsilon,$$

$$\|H^*(t) - Lt^*\| < \varepsilon t^*,$$

and

$$\|H^*(t) - Lt\| = \|H(t^*) - Lt^* + Lt^* - Lt\| \leq \|H(t^*) - Lt^*\| + \|L\|(t^* - t) < \varepsilon t^* + \|L\|\mu(\rho(t^*)).$$

Now, if t is such that $\rho(t^*) \geq T$, we obtain

$$\begin{aligned} \left\| \frac{H^*(t)}{t} - L \right\| &< \frac{\varepsilon t^*}{t} + \frac{\|L\|\mu(\rho(t^*))}{t} \leq \frac{\varepsilon t^*}{\rho(t^*)} + \frac{\|L\|\mu(\rho(t^*))}{\rho(t^*)} \\ &\leq \frac{\varepsilon\rho(t^*) + \mu(\rho(t^*))}{\rho(t^*)} + \frac{\|L\|\mu(\rho(t^*))}{\rho(t^*)} < \varepsilon + \varepsilon + \|L\|\varepsilon = \varepsilon(2 + \|L\|) \end{aligned}$$

and the conclusion follows easily. □

Combining the last lemma with Theorem 5.4, we obtain the following corollary.

Corollary 5.6. *If $\sup \mathbb{T} = \infty$, $\lim_{t\to\infty} \mu(t)/t = 0$, and $h : [t_0, \infty)_{\mathbb{T}} \to \mathbb{R}^n$ is Δ-integrable on every compact subinterval of $[t_0, \infty)_{\mathbb{T}}$, then*

$$\lim_{T\to\infty} \frac{1}{T} \int_{t_0}^{t_0+T} h^*(s)\,\mathrm{d}g(s) = \lim_{T\to\infty,\, t_0+T\in\mathbb{T}} \frac{1}{T} \int_{t_0}^{t_0+T} h(s)\,\Delta s,$$

provided the limit on the right-hand side exists.

Proof. Let $H(u) = \int_{t_0}^u h(s)\,\Delta s$, $u \in [t_0, \infty)_{\mathbb{T}}$. Then

$$\lim_{T\to\infty} \frac{1}{T} \int_{t_0}^{t_0+T} h^*(s)\,\mathrm{d}g(s) = \lim_{T\to\infty} \frac{H^*(t_0+T)}{T} = \lim_{T\to\infty} \frac{H^*(t_0+T)}{t_0+T}\frac{t_0+T}{T} = \lim_{T\to\infty} \frac{H^*(t_0+T)}{t_0+T}$$

$$= \lim_{u\to\infty} \frac{H^*(u)}{u} = \lim_{u\to\infty,\ u\in\mathbb{T}} \frac{H(u)}{u} = \lim_{T\to\infty,\ t_0+T\in\mathbb{T}} \frac{H(t_0+T)}{t_0+T} = \lim_{T\to\infty,\ t_0+T\in\mathbb{T}} \frac{H(t_0+T)}{T}\frac{T}{t_0+T}$$

$$= \lim_{T\to\infty,\ t_0+T\in\mathbb{T}} \frac{H(t_0+T)}{T} = \lim_{T\to\infty,\ t_0+T\in\mathbb{T}} \frac{1}{T}\int_{t_0}^{t_0+T} h(s)\,\Delta s. \qquad \square$$

In the context of our averaging theorem, we see that if $\lim_{t\to\infty} \mu(t)/t = 0$, then

$$f^0(x) = \lim_{T\to\infty} \frac{1}{T}\int_{t_0}^{t_0+T} f^*(x,s)\,\mathrm{d}g(s) = \lim_{T\to\infty,\ t_0+T\in\mathbb{T}} \frac{1}{T}\int_{t_0}^{t_0+T} f(x,s)\,\Delta s,$$

provided the right-hand side exists. This justifies the name "averaging theorem", since we are calculating the average of $t \mapsto f(x,t)$ over $[t_0,\infty)_{\mathbb{T}}$.

The condition $\lim_{t\to\infty} \mu(t)/t = 0$ means that the time scale graininess grows slower than linearly as $t \to \infty$. (In particular, every time scale with bounded graininess satisfies this condition.) Without this condition, the corollary is no longer true. To see this, assume that $\mu(t)/t > \varepsilon_0$ for arbitrarily large values t, and consider a function $h : [t_0,\infty)_{\mathbb{T}} \to \mathbb{R}$ which is identically equal to $L > 0$. If $t_0 + T \in \mathbb{T}$, we have $\frac{1}{T}\int_{t_0}^{t_0+T} h(s)\,\Delta s = L$. It follows from the properties of the Kurzweil-Stieltjes integral that the function $T \mapsto \frac{1}{T}\int_{t_0}^{t_0+T} h^*(s)\,\mathrm{d}g(s)$ has a jump of size $\frac{1}{T}L\mu(T)$ at every $T \in [t_0,\infty)_{\mathbb{T}}$, which exceeds $L\varepsilon_0$ for arbitrarily large values T. Thus the second limit does not exist.

Example 5.7. Consider the time scale $\mathbb{T} = \mathbb{Z}$ and the linear dynamic equation

$$x^{\Delta}(t) = \varepsilon\left(-1 + \frac{1}{1+t}\right)x(t),\ \ t \in \{0,1,2,\ldots\},\ \ x(0) = x_0.$$

Given an arbitrary $p \in \mathbb{R}$ and $r > 0$, the right-hand side $f(x,t) = \left(-1 + \frac{1}{1+t}\right)x$ is rd-continuous, bounded and Lipschitz-continuous in $B_r(p) \times [0,\infty)_{\mathbb{T}}$ (with Lipschitz-constant $l = 1$). Since $\lim_{t\to\infty} \mu(t)/t = 0$, we calculate

$$f^0(x,t) = \lim_{T\to\infty}\left(\frac{1}{T}\int_0^T \left(-1 + \frac{1}{1+t}\right)\Delta t\right)x = \left(\lim_{T\to\infty} \frac{-T + \sum_{i=1}^T \frac{1}{i}}{T}\right)x = -x.$$

According to the averaging theorem, the solutions of the original dynamic equation are well approximated by solutions of the averaged differential equation

$$y'(t) = -\varepsilon y(t),\ \ t \in [0,\infty),\ \ y(0) = x_0.$$

Note that we have deliberately chosen a simple example where both solutions can be found analytically:

$$\begin{aligned} x(t) &= e_{\varepsilon\left(-1+\frac{1}{1+t}\right)}(t,0)x_0 = \prod_{k=0}^{t-1}\left(1 - \varepsilon + \frac{\varepsilon}{1+k}\right)x_0 = \prod_{k=1}^{t}\left(1 - \varepsilon + \frac{\varepsilon}{k}\right)x_0 \\ y(t) &= e^{-\varepsilon t}x_0 \end{aligned}$$

6 Conclusion

The statement as well as the proof of the periodic averaging theorem for dynamic equations on time scales are similar to the classical case $\mathbb{T} = \mathbb{R}$. Generalizing the proof of the nonperiodic averaging theorem turns out to be more difficult and thus we have followed a different path via generalized ordinary differential equations. The disadvantage is that our result applies only to time scales with $\limsup_{t\to\infty} \mu(t)/t < \infty$, i.e. the graininess cannot grow faster than linearly as $t \to \infty$. The question whether there is an averaging theorem applicable to all time scales remains open.

Note also that in the periodic case, the averaged equation is a dynamic equation, while in the nonperiodic case, the averaged equation is an ordinary differential equation. (See also the paper [10], which contains a periodic averaging theorem where the averaged equation is an ordinary differential equation.) Both approximations have their merits: ordinary differential equations are often easier to deal with, but from an aesthetic point of view, it seems more natural to approximate a dynamic equation by a dynamic equation again. Thus the second open question is whether there exists a nonperiodic averaging theorem where the averaged equation is a dynamic equation defined on the same time scale as the original equation.

Acknowledgment. The author thanks the anonymous referee whose suggestions helped to improve this paper.

References

[1] E. P. Belan, *Averaging in the theory of finite-difference equations*, Ukrainian Mathematical Journal, Volume 19, Number 3, 319–323 (1967).

[2] M. Bohner, *Calculus of variations on time scales*, Dynamic Systems and Applications 13 (2004), 339–349.

[3] M. Bohner and G. Sh. Guseinov, *Partial differentiation on time scales*, Dynamic Systems and Applications 13 (2004), 351–379.

[4] M. Bohner and A. Peterson, *Dynamic Equations on Time Scales: An Introduction with Applications*, Birkhäuser, Boston, 2001.

[5] M. Bohner and A. Peterson, *Advances in Dynamic Equations on Time Scales*, Birkhäuser, Boston, 2003.

[6] H. S. Dumas, J. A. Ellison, and M. Vogt, *First-Order Averaging Principles for Maps with Applications to Accelerator Beam Dynamics*, SIAM J. Appl. Dyn. Syst. 3, No. 4, 409–432 (2004).

[7] R. Hilscher, V. Zeidan, and W. Kratz, *Differentiation of Solutions of Dynamic Equations on Time Scales with Respect to Parameters.* Advances in Dynamical Systems and Applications, Volume 4, Number 1, 35–54 (2009).

[8] S. G. Krantz, H. R. Parks, *The Implicit Function Theorem. History, Theory, and Applications*, Birkhäuser, 2002.

[9] J. Kurzweil, *Generalized ordinary differential equations and continuous dependence on a parameter*, Czech. Math. J. 7 (82), 418–449 (1957).

[10] J. G. Mesquita, A. Slavík, *Periodic averaging theorems for various types of equations*, J. Math. Anal. Appl. (2011), doi: 10.1016/j.jmaa.2011.09.038.

[11] C. C. Pugh, *Real Mathematical Analysis*, Springer, 2002.

[12] Š. Schwabik, *Generalized Ordinary Differential Equations*, World Scientific, Singapore, 1992.

[13] A. Slavík, *Dynamic equations on time scales and generalized ordinary differential equations*, J. Math. Anal. Appl. 385 (2012), 534–550.

[14] A. Slavík, *Product integration on time scales*, Dynamic Systems and Applications 19 (2010), 97–112.

[15] J. A. Sanders and F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems*, Springer-Verlag, New York, 1985.

[16] J. A. Sanders, F. Verhulst, and J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems* (2nd edition), Springer, New York, 2007.

[17] F. Verhulst, *Nonlinear Differential Equations and Dynamical Systems* (2nd edition), Springer, 2000.