1 Secular terms

Consider the undamped harmonic oscillator with time-dependent forcing $f(t)$,

$$\ddot{x}(t) + \omega_0^2 x(t) = f(t), \qquad x(0) = 0, \qquad \dot{x}(0) = 0. \tag{1}$$

The general solution for $x(t)$ is

$$x(t) = \frac{\sin(\omega_0 t)}{\omega_0} \int_0^t d\tilde{t} \cos(\omega_0 \tilde{t}) f(\tilde{t}) - \frac{\cos(\omega_0 t)}{\omega_0} \int_0^t d\tilde{t} \sin(\omega_0 \tilde{t}) f(\tilde{t}). \tag{2}$$

We assume that $f(t)$ is a periodic function of t with a period $T = 2\pi/\Omega$, $f(t) = f(t+T)$. Because of its periodicity, $f(t)$ can be expanded in a Fourier series,

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \sin(n\Omega t + \varphi_n). \tag{3}$$

Thus the solution Eq. (2) is an infinite sum,

$$x(t) = \sum_{n=0}^{\infty} x_n(t). \tag{4}$$

The individual summands are, for $n > 0$,

$$\begin{aligned} x_n(t) = {} & \frac{\omega_0 a_n \sin(\varphi_n)}{\omega_0 (n\Omega - \omega_0)(n\Omega + \omega_0)} \left(\cos(t\omega_0) - \cos(n\Omega t)\right) \\ & + \frac{a_n \cos(\varphi_n)}{\omega_0 (n\Omega - \omega_0)(n\Omega + \omega_0)} \left(n\Omega \sin(\omega_0 t) - \omega_0 \sin(n\Omega t)\right). \end{aligned} \tag{5}$$

As long as $\omega_0 \neq \Omega$, $x_n(t)$ is periodic in t for all $n \in \mathbb{N}$ and thus bounded for all times. The situation changes if the eigenfrequency ω_0 is an integer multiple of the forcing frequency Ω,

$$\Omega = m\omega_0, \; m \in \mathbb{N}. \tag{6}$$

The function $x_n(t)$ becomes

$$\begin{aligned} x_n(t) = {} & \frac{m^2 a_n \sin(\varphi_n)}{\omega_0^2 (m-n)(m+n)} \left(\cos\left(\frac{n}{m}\omega_0 t\right) - \cos(\omega_0 t)\right) \\ & + \frac{m a_n \cos(\varphi_n)}{\omega_0^2 (m-n)(m+n)} \left(m \sin\left(\frac{n}{m}\omega_0 t\right) - n \sin(\omega_0 t)\right). \end{aligned} \tag{7}$$

Note that the denominator as well as the numerator vanish for $n = m$. Taking the limit $n \to m$ and applying L'Hôspital's rule yields a term which is linear in t,

$$x_m(t) = \frac{a_m}{2\omega_0^2} \left(\cos(\varphi_m) \sin(\omega_0 t) - t\omega_0 \cos(\varphi_m + \omega_0 t)\right). \tag{8}$$

This term is called a *secular term* because it diverges for $t \to \infty$,

$$\lim_{t\to\infty} x_m(t) = \infty. \tag{9}$$

In conclusion, periodically forcing an undamped harmonic oscillator with a forcing frequency Ω identical to an integer multiple of the eigenfrequency ω_0 of the harmonic oscillator leads to an oscillation with an amplitude increasing linearly in time. Such secular terms diverge for large times. The exact form of the periodic forcing function $f(t)$ is not important.

2 Multiple scale perturbation expansion and averaging

Books which treat multiple scale perturbation expansion are [1] and [4]. For more information about averaging see the book [6]. Scholarpedia (http://www.scholarpedia.org) has articles about both topics written by active researchers in the field, see [2] and [5].
Consider the nonlinear system

$$\dot{\boldsymbol{x}}(t) = \epsilon \boldsymbol{f}(\boldsymbol{x}(t), t), \tag{10}$$

with initial condition

$$\boldsymbol{x}(t_0) = \boldsymbol{x}_0 + \epsilon \boldsymbol{x}_1. \tag{11}$$

The small parameter $0 \leq \epsilon \ll 1$ is used for a perturbation expansion. A regular perturbation expansion, i.e., a perturbation expansion in form of a Taylor series, may lead to so-called *secular terms*. Secular terms are terms which diverge as $t \to \infty$. Usually this does not reflect the true behavior of the system but arises as an artefact of the regular perturbation expansion. A *multiple scale perturbation* expansion may be able to avoid secular terms. Additional time scales are introduced to resolve phenomena on different time scales such as e.g. oscillations on a fast time scale and deformations of the periodic orbit on a slow time scale. In Eq. (10), an additional time $T = \epsilon(t - t_0)$ is introduced. The time derivative is substituted as

$$\frac{d}{dt} \to \frac{\partial}{\partial t} + \epsilon \frac{\partial}{\partial T}, \tag{12}$$

and the solution

$$\boldsymbol{x}(t) = \boldsymbol{x}(t, \epsilon(t - t_0)) = \boldsymbol{x}(t, T) \tag{13}$$

is assumed to depend on both time scales. The dynamical system Eq. (10) becomes

$$\frac{\partial}{\partial t}\boldsymbol{x}(t,T) + \epsilon \frac{\partial}{\partial T}\boldsymbol{x}(t,T) = \epsilon \boldsymbol{f}(\boldsymbol{x}(t,T), t). \tag{14}$$

From now on, the times t and T are assumed to be independent variables and Eq. (14) is considered as a *partial differential equation*. Only at the end of all computations the solution $\boldsymbol{x}$ is reverted back to the original single time scale t. The ansatz for the solution is an expansion of the form

$$\boldsymbol{x}(t,T) = \boldsymbol{X}_0(t,T) + \epsilon \boldsymbol{X}_1(t,T) + \mathcal{O}\left(\epsilon^2\right). \tag{15}$$

Equation (14) becomes, up to $\mathcal{O}\left(\epsilon^2\right)$,

$$\frac{\partial}{\partial t}\boldsymbol{X}_0(t,T) + \epsilon\frac{\partial}{\partial t}\boldsymbol{X}_1(t,T) + \epsilon\frac{\partial}{\partial T}\boldsymbol{X}_0(t,T) = \epsilon \boldsymbol{f}\left(\boldsymbol{X}_0(t,T),t\right) + \mathcal{O}\left(\epsilon^2\right). \tag{16}$$

Collecting terms in orders of ϵ yields in $\mathcal{O}(1)$

$$\frac{\partial}{\partial t}\boldsymbol{X}_0(t,T) = 0. \tag{17}$$

Its solution $\boldsymbol{X}_0(t,T)$ is constant in t but may nevertheless depend on T,

$$\boldsymbol{X}_0(t,T) = \boldsymbol{X}_0(t_0,T) = \hat{\boldsymbol{X}}_0(T). \tag{18}$$

The equation obtained in $\mathcal{O}(\epsilon)$ is

$$\frac{\partial}{\partial t}\boldsymbol{X}_1(t,T) + \hat{\boldsymbol{X}}_0'(T) = \boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right). \tag{19}$$

To apply averaging, we assume that $\boldsymbol{X}_1(t,T)$ is periodic in t with period T_0,

$$\boldsymbol{X}_1(t,T) = \boldsymbol{X}_1(t+T_0,T), \tag{20}$$

such that

$$\left\langle \frac{\partial}{\partial t}\boldsymbol{X}_1(t,T)\right\rangle = \frac{1}{T_0}\left(\boldsymbol{X}_1(T_0,T) - \boldsymbol{X}_1(0,T)\right) = 0. \tag{21}$$

The period T_0 is usually determined from the context, as e.g. the periodicity of $\boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right)$ in t,

$$\boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right) = \boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t+T_0\right). \tag{22}$$

The angular brackets denote averaging over one temporal period,

$$\langle \boldsymbol{y}(t)\rangle = \frac{1}{T_0}\int_0^{T_0} \mathrm{d}t \boldsymbol{y}(t). \tag{23}$$

Averaging Eq. (19) over one period yields a differential equation for $\hat{\boldsymbol{X}}_0(T)$ which is independent of the fast time scale t,

$$\hat{\boldsymbol{X}}_0'(T) = \left\langle \boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right)\right\rangle = \frac{1}{T_0}\int_0^{T_0}\mathrm{d}t \boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right). \tag{24}$$

Finally, using Eq. (24) in Eq. (19) yields a differential equation for $\boldsymbol{X}_1$

$$\frac{\partial}{\partial t}\boldsymbol{X}_1(t,T) = \boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right) - \left\langle \boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right)\right\rangle \tag{25}$$

with solution

$$\boldsymbol{X}_1(t,T) = \boldsymbol{X}_1(t_0,T) + \int_{t_0}^{t} \mathrm{d}\tilde{t}\boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),\tilde{t}\right) - (t-t_0)\left\langle \boldsymbol{f}\left(\hat{\boldsymbol{X}}_0(T),t\right)\right\rangle. \tag{26}$$

The approximate solution up to $\mathcal{O}\left(\epsilon^2\right)$ is given by

$$\boldsymbol{x}(t,T) = \hat{\boldsymbol{X}}_0(T) + \epsilon\boldsymbol{X}_1(t,T), \tag{27}$$

or, reverting back to the original single time scale t,

$$\boldsymbol{x}(t) = \boldsymbol{x}(t,\epsilon(t-t_0)) = \hat{\boldsymbol{X}}_0(\epsilon(t-t_0)) + \epsilon\boldsymbol{X}_1(t,\epsilon(t-t_0)). \tag{28}$$

Finally, the initial conditions follow from Eq. (11) as

$$\hat{\boldsymbol{X}}_0(0) = \boldsymbol{x}_0, \qquad \boldsymbol{X}_1(t_0,0) = \boldsymbol{x}_1. \tag{29}$$

3 Floquet theory

3.1 State transition matrix and fundamental state transition matrix

See the book [7], page 88, for an alternative proof of Floquet's theorem. The book is a readable introduction about the theory of dynamical systems and differential equations from a more mathematical point of view.
We consider the linear dynamical system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{\mathcal{A}}(t)\boldsymbol{x}(t), \qquad \boldsymbol{x}(t_0) = \boldsymbol{x}_0, \tag{30}$$

with time-dependent *state matrix* $\boldsymbol{\mathcal{A}}(t)$. A *state transition matrix* $\boldsymbol{\mathcal{U}}(t)$ satisfies

$$\dot{\boldsymbol{\mathcal{U}}}(t) = \boldsymbol{\mathcal{A}}(t)\boldsymbol{\mathcal{U}}(t). \tag{31}$$

The solution $\boldsymbol{\mathcal{U}}(t)$ to Eq. (31) is not unique because no initial condition is specified. The solution for $\boldsymbol{x}(t)$ is given by

$$\boldsymbol{x}(t) = \boldsymbol{\mathcal{U}}(t)\boldsymbol{\mathcal{U}}^{-1}(t_0)\boldsymbol{x}_0, \tag{32}$$

which is easily proven,

$$\dot{\boldsymbol{x}}(t) = \dot{\boldsymbol{\mathcal{U}}}(t)\boldsymbol{\mathcal{U}}^{-1}(t_0)\boldsymbol{x}_0 = \boldsymbol{\mathcal{A}}(t)\boldsymbol{\mathcal{U}}(t)\boldsymbol{\mathcal{U}}^{-1}(t_0)\boldsymbol{x}_0 = \boldsymbol{\mathcal{A}}(t)\boldsymbol{x}(t), \tag{33}$$

$$\boldsymbol{x}(t_0) = \boldsymbol{\mathcal{U}}(t_0)\boldsymbol{\mathcal{U}}^{-1}(t_0)\boldsymbol{x}_0 = \boldsymbol{x}_0. \tag{34}$$

The state transition matrix

$$\boldsymbol{\Phi}\left(t,t_0\right)=\boldsymbol{\mathcal{U}}\left(t\right)\boldsymbol{\mathcal{U}}^{-1}\left(t_0\right) \tag{35}$$

is called the *fundamental state transition matrix* and satisfies

$$\partial_t\boldsymbol{\Phi}\left(t,t_0\right)=\boldsymbol{\mathcal{A}}\left(t\right)\boldsymbol{\Phi}\left(t,t_0\right), \qquad \boldsymbol{\Phi}\left(t_0,t_0\right)=\mathbf{1}. \tag{36}$$

While there is a continuous family of state transition matrices $\boldsymbol{\mathcal{U}}\left(t\right)$ parametrized by their initial conditions $\boldsymbol{\mathcal{U}}\left(t_0\right)$, the fundamental state transition matrix $\boldsymbol{\Phi}\left(t,t_0\right)$ is unique.
Consider two state transition matrices $\boldsymbol{\mathcal{U}}_1\left(t\right)$ and $\boldsymbol{\mathcal{U}}_2\left(t\right)$. Because of the uniqueness of $\boldsymbol{\Phi}\left(t,t_0\right)$, we obtain

$$\boldsymbol{\Phi}\left(t,t_0\right)=\boldsymbol{\mathcal{U}}_1\left(t\right)\boldsymbol{\mathcal{U}}_1^{-1}\left(t_0\right)=\boldsymbol{\mathcal{U}}_2\left(t\right)\boldsymbol{\mathcal{U}}_2^{-1}\left(t_0\right).$$

Thus, two state transition matrices do only differ by a constant multiplicative matrix $\boldsymbol{\mathcal{C}}$,

$$\boldsymbol{\mathcal{U}}_1\left(t\right)=\boldsymbol{\mathcal{U}}_2\left(t\right)\boldsymbol{\mathcal{C}}, \tag{37}$$

with

$$\boldsymbol{\mathcal{C}}=\boldsymbol{\mathcal{U}}_2^{-1}\left(t_0\right)\boldsymbol{\mathcal{U}}_1\left(t_0\right)=\text{const.} \tag{38}$$

3.2 Periodic state matrix

Floquet theory is concerned with linear dynamical systems with periodic state matrices. In particular, Gaston Floquet showed in 1883 [3] that a state transformation exists which yields a linear dynamical system with a constant state matrix. In solid-state physics, an analogous result is known as Bloch's theorem.
Consider Eq. (30) with a state matrix periodic in t with period T,

$$\boldsymbol{\mathcal{A}}\left(t+T\right)=\boldsymbol{\mathcal{A}}\left(t\right). \tag{39}$$

If $\boldsymbol{\mathcal{U}}\left(t\right)$ is a state transition transition matrix, then $\boldsymbol{\mathcal{U}}\left(t+T\right)$ is a state transition matrix as well,

$$\frac{d}{dt}\boldsymbol{\mathcal{U}}\left(t+T\right)=\boldsymbol{\mathcal{A}}\left(t+T\right)\boldsymbol{\mathcal{U}}\left(t+T\right)=\boldsymbol{\mathcal{A}}\left(t\right)\boldsymbol{\mathcal{U}}\left(t+T\right). \tag{40}$$

Thus, $\boldsymbol{\mathcal{U}}\left(t+T\right)$ and $\boldsymbol{\mathcal{U}}\left(t\right)$ are both state transition matrices. As was shown in Eq. (37), both can only differ by a multiplicative constant matrix $\boldsymbol{\mathcal{C}}$,

$$\boldsymbol{\mathcal{U}}\left(t+T\right)=\boldsymbol{\mathcal{U}}\left(t\right)\boldsymbol{\mathcal{C}}=\boldsymbol{\mathcal{U}}\left(t\right)\exp\left(\boldsymbol{\mathcal{B}}T\right). \tag{41}$$

We defined a new matrix $\boldsymbol{\mathcal{B}}$ by the relation

$$\boldsymbol{\mathcal{C}}=\exp\left(\boldsymbol{\mathcal{B}}T\right). \tag{42}$$

The matrix $\boldsymbol{\mathcal{B}}$ is constant and, in general, complex. The matrix $\boldsymbol{\mathcal{P}}(t)$ defined by

$$\boldsymbol{\mathcal{P}}(t) = \boldsymbol{\mathcal{U}}(t)\exp(-\boldsymbol{\mathcal{B}}t)$$

is periodic in t with period T,

$$\begin{aligned}\boldsymbol{\mathcal{P}}(t+T) &= \boldsymbol{\mathcal{U}}(t+T)\exp(-\boldsymbol{\mathcal{B}}T)\exp(-\boldsymbol{\mathcal{B}}t)\\ &= \boldsymbol{\mathcal{U}}(t)\exp(-\boldsymbol{\mathcal{B}}t) = \boldsymbol{\mathcal{P}}(t).\end{aligned} \tag{43}$$

Furthermore, $\boldsymbol{\mathcal{P}}(t)$ is invertible with an inverse given by

$$\boldsymbol{\mathcal{P}}^{-1}(t) = \exp(\boldsymbol{\mathcal{B}}t)\,\boldsymbol{\mathcal{U}}^{-1}(t). \tag{44}$$

The matrix $\boldsymbol{\mathcal{P}}(t)$ transforms the linear dynamical system Eq. (30) with periodic state matrix Eq. (39) to a system with a constant state matrix. Let the new state vector be

$$\boldsymbol{y}(t) = \boldsymbol{\mathcal{P}}^{-1}(t)\,\boldsymbol{x}(t), \tag{45}$$

and the inverse transformation is

$$\boldsymbol{x}(t) = \boldsymbol{\mathcal{P}}(t)\,\boldsymbol{y}(t). \tag{46}$$

Applying the time derivative onto $\boldsymbol{x}(t)$ yields

$$\begin{aligned}\dot{\boldsymbol{x}}(t) &= \dot{\boldsymbol{\mathcal{P}}}(t)\,\boldsymbol{y}(t) + \boldsymbol{\mathcal{P}}(t)\,\dot{\boldsymbol{y}}(t)\\ &= \dot{\boldsymbol{\mathcal{U}}}(t)\exp(-\boldsymbol{\mathcal{B}}t)\,\boldsymbol{y}(t) - \boldsymbol{\mathcal{U}}(t)\exp(-\boldsymbol{\mathcal{B}}t)\,\boldsymbol{\mathcal{B}}\boldsymbol{y}(t) + \boldsymbol{\mathcal{P}}(t)\,\dot{\boldsymbol{y}}(t)\\ &= \boldsymbol{\mathcal{A}}(t)\,\boldsymbol{\mathcal{U}}(t)\exp(-\boldsymbol{\mathcal{B}}t)\,\boldsymbol{y}(t) - \boldsymbol{\mathcal{P}}(t)\,\boldsymbol{\mathcal{B}}\boldsymbol{y}(t) + \boldsymbol{\mathcal{P}}(t)\,\dot{\boldsymbol{y}}(t)\\ &= \boldsymbol{\mathcal{A}}(t)\,\boldsymbol{\mathcal{P}}(t)\,\boldsymbol{y}(t) - \boldsymbol{\mathcal{P}}(t)\,\boldsymbol{\mathcal{B}}\boldsymbol{y}(t) + \boldsymbol{\mathcal{P}}(t)\,\dot{\boldsymbol{y}}(t)\\ &= \boldsymbol{\mathcal{A}}(t)\,\boldsymbol{x}(t) - \boldsymbol{\mathcal{P}}(t)\,\boldsymbol{\mathcal{B}}\boldsymbol{y}(t) + \boldsymbol{\mathcal{P}}(t)\,\dot{\boldsymbol{y}}(t).\end{aligned} \tag{47}$$

Exploiting Eq. (30) finally yields a linear dynamical system with constant state matrix $\boldsymbol{\mathcal{B}}$,

$$\dot{\boldsymbol{y}}(t) = \boldsymbol{\mathcal{B}}\boldsymbol{y}(t). \tag{48}$$

Example. *Stability of a limit cycle.*
Consider the nonlinear dynamical system

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{R}(\boldsymbol{x}(t)) \tag{49}$$

with limit cycle solution with period T,

$$\boldsymbol{x}(t) = \boldsymbol{X}_0(t) = \boldsymbol{X}_0(t+T). \tag{50}$$

To investigate the linear stability of the limit cycle $\boldsymbol{X}_0(t)$, we study the approximate solution to Eq. (49) with a perturbed initial condition lying slightly off the limit cycle,

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{R}(\boldsymbol{x}(t)), \qquad \boldsymbol{x}(t_0) = \boldsymbol{X}_0(t_0) + \epsilon\boldsymbol{x}_1. \tag{51}$$

With the ansatz

$$\boldsymbol{x}(t) = \boldsymbol{X}_0(t) + \epsilon \boldsymbol{X}_1(t), \tag{52}$$

we obtain in first order of ϵ a linear dynamical system with periodic state matrix $\boldsymbol{\mathcal{A}}(t)$,

$$\dot{\boldsymbol{X}}_1(t) = \boldsymbol{\mathcal{A}}(t)\boldsymbol{X}_1(t), \qquad \boldsymbol{X}_1(t_0) = \boldsymbol{x}_1, \tag{53}$$

with

$$\boldsymbol{\mathcal{A}}(t) = \nabla_{\boldsymbol{x}} \boldsymbol{R}(\boldsymbol{X}_0(t)) = \nabla_{\boldsymbol{x}} \boldsymbol{R}(\boldsymbol{X}_0(t+T)) = \boldsymbol{\mathcal{A}}(t+T). \tag{54}$$

With the help of the transform defined by Floquet, we can express the solution $\boldsymbol{X}_1(t)$ as

$$\begin{aligned}\boldsymbol{X}_1(t) &= \boldsymbol{\mathcal{P}}(t)\boldsymbol{y}(t) = \boldsymbol{\mathcal{P}}(t)\exp(\boldsymbol{\mathcal{B}}(t-t_0))\boldsymbol{y}_0 \\ &= \boldsymbol{\mathcal{P}}(t)\exp(\boldsymbol{\mathcal{B}}(t-t_0))\boldsymbol{\mathcal{P}}^{-1}(t_0)\boldsymbol{x}_1.\end{aligned} \tag{55}$$

Assuming that $\boldsymbol{\mathcal{B}}$ can be diagonalized, we can introduce the matrix $\boldsymbol{\mathcal{T}}$ which diagonalizes $\boldsymbol{\mathcal{B}}$ as

$$\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{T}}\boldsymbol{\mathcal{B}}_D\boldsymbol{\mathcal{T}}^{-1}, \tag{56}$$

and $\boldsymbol{\mathcal{B}}_D$ is a diagonal matrix with the eigenvalues of $\boldsymbol{\mathcal{B}}$ on its diagonal. The solution $\boldsymbol{X}_1(t)$ can be written as

$$\boldsymbol{X}_1(t) = \boldsymbol{\mathcal{P}}(t)\boldsymbol{\mathcal{T}}\exp(\boldsymbol{\mathcal{B}}_D(t-t_0))\boldsymbol{\mathcal{T}}^{-1}\boldsymbol{\mathcal{P}}^{-1}(t_0)\boldsymbol{x}_1.$$

Because the matrix $\boldsymbol{\mathcal{P}}(t)$ is periodic it must be bounded. Thus, the asymptotic behavior of $\boldsymbol{X}_1(t)$ for large times can be read off Eq. (55). If the real parts $\Re(\lambda_i)$ of all eigenvalues λ_i, $i \in \{1, \ldots, n\}$, of $\boldsymbol{\mathcal{B}}$ are strictly negative, then

$$\lim_{t\to\infty} \boldsymbol{X}_1(t) = \boldsymbol{0}, \tag{57}$$

and the limit cycle solution $\boldsymbol{X}_0(t)$ is linearly stable.

References

[1] Carl M Bender and Steven A Orszag. *Advanced Mathematical Methods for Scientists and Engineers I*. Springer Science & Business Media, 1999.

[2] C. C. Chow. Multiple scale analysis. *Scholarpedia*, 2(10):1617, 2007.

[3] Gaston Floquet. Sur les équations différentielles linéaires à coefficients périodiques. In *Annales scientifiques de l'École normale supérieure*, volume 12, pages 47–88, 1883.

[4] J.K. Kevorkian and J.D. Cole. *Multiple Scale and Singular Perturbation Methods*. Springer, 1996.

[5] J. A. Sanders. Averaging. *Scholarpedia*, 1(11):1760, 2006.

[6] Jan A Sanders, Ferdinand Verhulst, and James A Murdock. *Averaging methods in nonlinear dynamical systems*, volume 2. Springer, 2007.

[7] Gerald Teschl. *Ordinary Differential Equations and Dynamical Systems*. American Mathematical Society, 2012.